

# Mathematical Foundations of Proof-Theoretic Ordinals

## 1.1 Recursive Ordinals and Ordinal Arithmetic

**Definition:** A **recursive ordinal** is an ordinal  $\alpha$  such that there exists a recursive well-ordering of  $\mathbb{N}$  of order type  $\alpha$ .

Recursive ordinals form a proper initial segment of the class of all ordinals, with supremum the Church-Kleene ordinal  $\omega_1^{\text{CK}}$ .

### Fundamental Operations

**Ordinal Addition:**  $\alpha + \beta$

- Concatenation of orderings
- Non-commutative:  $1 + \omega = \omega \neq \omega + 1$

**Ordinal Multiplication:**  $\alpha \cdot \beta$

- $\beta$  copies of  $\alpha$  in succession
- Non-commutative:  $2 \cdot \omega = \omega \neq \omega \cdot 2$

**Ordinal Exponentiation:**  $\omega^\alpha$

- Limit case:  $\omega^\alpha = \sup\{\omega^\beta : \beta < \alpha\}$  if  $\alpha$  is limit
- Successor case:  $\omega^{(\alpha+1)} = \omega^\alpha \cdot \omega$

## 1.2 Cantor Normal Form

**Theorem (Cantor):** Every ordinal  $\alpha > 0$  can be uniquely represented as:

$$\alpha = \omega^{\beta_1} \cdot c_1 + \omega^{\beta_2} \cdot c_2 + \dots + \omega^{\beta_\square} \cdot c_\square$$

where:

- $\beta_1 > \beta_2 > \dots > \beta_\square \geq 0$
- $c_i \in \mathbb{N} \setminus \{0\}$  for all  $i$
- Each  $\beta_i$  is itself in normal form

**Examples:**

- $42 = \omega^0 \cdot 42$
- $\omega^2 + \omega \cdot 3 + 5$
- $\omega^\omega = \omega^\omega$  (already in normal form)

- $\omega^{(\omega^2+1)} \cdot 2 + \omega^5 \cdot 7 + 3$

## 1.3 Fixed Points of the Exponential Function

Consider the function  $\varphi: \alpha \mapsto \omega^\alpha$ .

**Definition:** An ordinal  $\varepsilon$  is a **fixed point** if  $\omega^\varepsilon = \varepsilon$ .

**Construction of the  $\varepsilon$ -hierarchy:**

- $\varepsilon_0 = \sup\{\omega, \omega^\omega, \omega^{\omega^\omega}, \omega^{\omega^{\omega^\omega}}, \dots\}$
- $\varepsilon_{a+1} = \text{first } \varepsilon > \varepsilon_a \text{ such that } \omega^\varepsilon = \varepsilon$
- $\varepsilon_\omega = \sup\{\varepsilon_0, \varepsilon_1, \varepsilon_2, \dots\}$
- $\varepsilon_{\omega^2}(\varepsilon_0) = \dots$

**Proposition:** For every  $\alpha$ ,  $\varepsilon_\alpha$  is well-defined.

**Crucial characterization of  $\varepsilon_0$ :**

- It is the first ordinal not representable in Cantor normal form using only  $\omega$ ,  $+$ ,  $\cdot$ ,  $^\wedge$
- It is the minimum ordinal closed under exponentiation
- For every  $\alpha < \varepsilon_0$ , there exist  $\alpha_1, \dots, \alpha_n < \alpha$  such that  $\alpha = \omega^{\alpha_1} + \dots + \omega^{\alpha_n}$

## 1.4 Ordinal Notations

An **ordinal notation system** is a method for representing ordinals via finite objects (strings, trees, etc.).

**Cantor system for ordinals  $< \varepsilon_0$ :**

Grammar:

$T ::= 0 \mid \omega^T \mid T + T \mid c \cdot T \quad (c \in \mathbb{N} \setminus \{0\})$

With ordering relation  $<$  defined recursively.

**Property:** Every well-formed term denotes a unique ordinal  $< \varepsilon_0$ , and every ordinal  $< \varepsilon_0$  has a unique canonical representation.

## 1.5 The Veblen Hierarchy

To go beyond  $\varepsilon_0$ , Veblen (1908) introduced the hierarchy of  $\varphi$  functions.

**Veblen function  $\varphi_\alpha(\beta)$ :**

- $\varphi_0(\beta) = \omega^\beta$
- $\varphi_{a+1}(\beta) = \beta$ -th fixed point of  $\varphi_a$

- $\varphi_{\lambda}(\beta) = \beta$ -th common fixed point of all  $\varphi_{\alpha}$  for  $\alpha < \lambda$

**Examples:**

- $\varphi_1(0) = \varepsilon_0$
- $\varphi_1(1) = \varepsilon_1$
- $\varphi_1(\omega) = \varepsilon_{\omega}$
- $\varphi_2(0) =$  first fixed point of  $\alpha \mapsto \varepsilon_{\alpha}$
- $\Gamma_0 = \varphi_{\Gamma_0}(0) =$  first fixed point of  $\alpha \mapsto \varphi_{\alpha}(0)$

## 1.6 Importance for Proof Theory

Ordinals  $< \varepsilon_0$  are exactly those predicatively acceptable according to the Feferman-Schütte analysis. They represent the limit of:

1. Hilbert's extended finitism (Gentzen 1936)
2. Predicative mathematics (Weyl, Feferman)
3. Proof strength of Peano Arithmetic PA

The  $\varepsilon_0$  barrier marks the transition from finitistic methods to genuinely infinitistic methods in proof theory.

## 1.7 The Veblen Binary Function $\varphi(\alpha, \beta)$

To handle arbitrary indices uniformly:

**Definition:**

- $\varphi(0, \beta) := \omega^{\beta}$
- $\varphi(\alpha+1, \beta) := \beta$ -th fixed point of  $\gamma \mapsto \varphi(\alpha, \gamma)$
- $\varphi(\lambda, \beta)$  for limit  $\lambda := \beta$ -th ordinal that is a simultaneous fixed point of all  $\varphi(\alpha, \cdot)$  for  $\alpha < \lambda$

**Relation to previous notation:**

- $\varphi(n, \beta) = \varphi_n(\beta)$  for finite  $n$
- $\varphi(\omega, 0) = \min\{\beta : \varphi(n, \beta) = \beta \text{ for all } n < \omega\}$

**Fundamental properties:**

1. **Monotonicity:** If  $\alpha_1 \leq \alpha_2$  and  $\beta_1 \leq \beta_2$ , then  $\varphi(\alpha_1, \beta_1) \leq \varphi(\alpha_2, \beta_2)$
2. **Normality:**  $\varphi(\alpha, \cdot)$  is continuous for every  $\alpha$
3. **Fixed points:**  $\varphi(\alpha+1, \beta)$  is always a fixed point of  $\varphi(\alpha, \cdot)$
4. **Enumeration:**  $\varphi(\alpha, \cdot)$  enumerates exactly the common fixed points of all  $\varphi(\gamma, \cdot)$  for  $\gamma < \alpha$

## 1.8 Veblen Normal Form

**Theorem:** Every ordinal  $\alpha < \Gamma_0$  can be uniquely represented as:

$$\alpha = \varphi(\alpha_1, \beta_1) + \varphi(\alpha_2, \beta_2) + \dots + \varphi(\alpha_\square, \beta_\square)$$

where:

- $\varphi(\alpha_1, \beta_1) \geq \varphi(\alpha_2, \beta_2) \geq \dots \geq \varphi(\alpha_\square, \beta_\square)$
- Each  $\alpha_i, \beta_i < \alpha$  is itself in normal form
- Each term is principal:  $\varphi(\alpha_i, \beta_i)$  is not a sum of smaller ordinals

**Example representations:**

- $\varphi(1, 0) + \omega^2 + 5 = \varepsilon_0 + \omega^2 + 5$
- $\varphi(2, 3) \cdot \omega + \varphi(1, \varphi(1, 0)) + \varphi(0, \varphi(0, 7))$

## 1.9 The Feferman-Schütte Ordinal $\Gamma_0$

**Problem:** The notation  $\varphi(\alpha, \beta)$  requires  $\alpha$  and  $\beta$  to be already defined, but for large  $\alpha$  we want to use the  $\varphi$  notation itself.

**Solution:** Find the first fixed point of the function  $\alpha \mapsto \varphi(\alpha, 0)$ .

**Definition:**  $\Gamma_0 := \min\{\alpha : \varphi(\alpha, 0) = \alpha\}$

**Equivalently:**

- $\Gamma_0 = \sup\{\varphi(0, 0), \varphi(\varphi(0, 0), 0), \varphi(\varphi(\varphi(0, 0), 0), 0), \dots\}$
- $\Gamma_0$  is the first ordinal that "contains itself" in the  $\varphi$  notation

**Explicit construction:**

Define recursively:

- $\gamma_0 = 0$
- $\gamma_{\square+1} = \varphi(\gamma_\square, 0)$
- $\Gamma_0 = \sup\{\gamma_\square : \square \in \mathbb{N}\}$

**Initial sequence:**

- $\gamma_0 = 0$
- $\gamma_1 = \varphi(0, 0) = \omega^0 = 1$
- $\gamma_2 = \varphi(1, 0) = \varepsilon_0$
- $\gamma_3 = \varphi(\varepsilon_0, 0) = \zeta_0$  (first fixed point of  $\alpha \mapsto \varepsilon_\alpha$ )
- $\gamma_4 = \varphi(\zeta_0, 0)$
- ...

## 1.10 Extended Veblen Hierarchy

To go beyond  $\Gamma_0$ , Veblen introduced functions of arbitrary finite arity:

**n-ary Veblen function  $\varphi(\alpha_1, \alpha_2, \dots, \alpha_n, \beta)$ :**

- $\varphi(\beta) := \omega^\beta$
- $\varphi(\alpha_1, \dots, \alpha_n, \beta) := \beta$ -th common fixed point of all  $\varphi(\alpha_1, \dots, \alpha_{n-1}, \cdot)$  where  $\alpha_n < \alpha_n$

**Notable ordinals:**

- $\Gamma_0 = \varphi(1, 0, 0)$
- $\Gamma_1 = \varphi(1, 0, 1)$
- $\Gamma_{\Gamma_0} = \varphi(1, 0, \Gamma_0)$
- $\varphi(1, 1, 0) =$  first fixed point of  $\alpha \mapsto \varphi(1, 0, \alpha)$
- $\varphi(2, 0, 0) =$  first fixed point of  $\alpha \mapsto \varphi(1, \alpha, 0)$

**Limit of Veblen hierarchy:**

The **small Veblen ordinal** is:  $SVO = \sup\{\varphi(\dots), \varphi(\varphi(\dots), \dots), \dots\}$

where we consider all Veblen functions of finite arity.

## 1.11 Ordinals and Formal Systems

**Fundamental correspondences:**

| Ordinal                          | System                | Characteristic Principle            |
|----------------------------------|-----------------------|-------------------------------------|
| $\omega^\omega$                  | PRA                   | Primitive recursive induction       |
| $\varepsilon_0$                  | PA                    | First-order induction               |
| $\varphi(\omega, 0)$             | $ACA_0$               | Arithmetical comprehension          |
| $\Gamma_0$                       | $ATR_0$               | Arithmetical transfinite recursion  |
| $\psi_0(\varepsilon_{\Omega+1})$ | $\Pi^1_1\text{-}CA_0$ | $\Pi^1_1$ -comprehension            |
| $\psi(\Omega_\omega)$            | $ID_1$                | Inductive definitions iterated once |

## 1.12 The Bachmann-Howard Function $\psi$

To handle truly large ordinals (beyond SVO), a radically different notation is needed.

**Idea:** Use "non-constructible" (uncountable) ordinals as "markers" to collapse large ordinals.

**Uncountable ordinals:**

- $\Omega$  = first uncountable ordinal
- $\Omega_2$  = second uncountable ordinal

- ...

**Bachmann-Howard collapsing function:**

$\psi_\Omega(\alpha)$  := the  $\alpha$ -th ordinal constructible using only:

- 0, +, Veblen  $\varphi$
- ordinals  $< \Omega$
- previous results of  $\psi_\Omega$

**Extraordinary property:**  $\psi_\Omega(\varepsilon_-(\Omega+1)) = \text{ordinal of } \Pi^1_1\text{-CA}_0$

This function allows denoting ordinals vastly beyond the Veblen hierarchy, using  $\Omega$  as a "marker" that is then "collapsed away".

## 1.13 Ordinal Collapsing Technique

**General principle:** For every large ordinal  $\alpha$ :

1. Work in a universe with uncountable ordinals  $\Omega, \Omega_2, \dots$
2. Construct expressions using  $\alpha$  and these uncountable ordinals
3. "Collapse" the expression by removing references to the  $\Omega$ s

**Example:**

- $\psi(\Omega^\Omega)$  denotes an enormous ordinal, but "countable"
- It is far beyond anything in the finite Veblen hierarchy
- It is still recursive and can be used to measure proof strength

## 1.14 Connection with Proof Theory

These large ordinals measure the complexity of proofs in increasingly strong systems:

- The stronger the logical system
- The larger the ordinal needed for its analysis
- The more powerful normalization techniques required

Hilbert's program, even after Gödel, continues through these ordinal analyses.

## Part 2: Gentzen's Theorem (1936)

### 2.1 Historical Context

After Gödel's incompleteness theorems (1931), it became clear that PA cannot prove its own consistency. Gentzen demonstrated that by adding a minimal transfinite principle, one can prove  $\text{Con}(\text{PA})$ .

This result is foundational for modern proof theory and provides a precise measure of PA's logical strength.

## 2.2 Transfinite Induction: Precise Formulation

**Definition (Transfinite Induction for an ordinal  $\alpha$ ):**

For a predicate  $P$  and an ordinal  $\alpha$ , we define:

$$\text{TI}(\alpha, P): [\forall \beta < \alpha (\forall \gamma < \beta P(\gamma) \rightarrow P(\beta))] \rightarrow \forall \beta < \alpha P(\beta)$$

**Equivalently:** If  $P$  is progressive along  $\alpha$  (i.e.,  $P(\beta)$  holds whenever  $P$  holds for all smaller ordinals), then  $P$  holds everywhere below  $\alpha$ .

**Schema  $\text{TI}(\alpha)$ :** For every formula  $\phi(x)$  of the language, the axiom  $\text{TI}(\alpha, \lambda x.\phi(x))$

## 2.3 Representation of Ordinals in Arithmetic

To speak of  $\text{TI}(\alpha)$  in PA, we must encode ordinals as relations on  $\mathbb{N}$ .

**Definition:** A binary relation  $< \subseteq \mathbb{N}^2$  is a **well-ordering** if:

1.  $<$  is a linear ordering
2. Every non-empty subset has a  $<$ -minimal element

### Encoding $\epsilon_0$ in PA

We represent ordinals  $< \epsilon_0$  via terms in Cantor normal form, encoded as natural numbers. We define  $<_{\epsilon_0}$  as the ordering relation on these codes.

**Coding function:**  $\text{Cod}: \{\text{ordinals} < \epsilon_0\} \rightarrow \mathbb{N}$

$$0 \mapsto \langle 0 \rangle$$

$$\omega^\alpha \mapsto \langle 1, \text{Cod}(\alpha) \rangle$$

$$\alpha + \beta \mapsto \langle 2, \text{Cod}(\alpha), \text{Cod}(\beta) \rangle$$

...

Then we define in PA a formula  $\text{BO}(x)$  that says "x codes a well-ordering", thus obtaining a formula that "expresses" the property of being well-ordered.

## 2.4 Gentzen's Theorem: Precise Statement

**Theorem (Gentzen, 1936):**

(i)  $\text{PA} + \text{TI}(\epsilon_0) \vdash \text{Con}(\text{PA})$

(ii) PA does not prove  $\text{TI}(\epsilon_0)$

(iii) For every  $\alpha < \varepsilon_0$ ,  $PA \vdash TI(\alpha)$

**Corollary:**  $\varepsilon_0$  is exactly the proof-theoretic ordinal of PA, written  $|PA| = \varepsilon_0$ .

## 2.5 Proof Sketch of (i)

The proof proceeds in three steps:

### Step 1: Normalization in the Sequent Calculus

Gentzen introduces the sequent calculus LK and proves that:

- PA is interpretable in LK
- Every derivation in LK can be transformed into a cut-free derivation
- The length of the derivation may increase, but logical complexity decreases

### Step 2: Assignment of Ordinals to Derivations

To each cut-free derivation  $D$ , we assign an ordinal  $o(D) < \varepsilon_0$  that measures:

- The depth of the derivation
- The number and type of rules used
- The complexity of formulas

**Crucial property:** If  $D_1, \dots, D_n$  are the immediate premises of  $D$ , then:  $o(D_i) < o(D)$  for all  $i$

### Step 3: Transfinite Induction

Let  $P(\alpha) := \text{"There is no derivation of } \perp \text{ with ordinal } \alpha\text{"}$

- **Base:**  $P(0)$  is obvious
- **Inductive step:** If  $P(\beta)$  for all  $\beta < \alpha$ , then  $P(\alpha)$ 
  - Suppose there exists  $D$  with  $o(D) = \alpha$  deriving  $\perp$
  - Its premises have ordinals  $< \alpha$
  - By inductive hypothesis, they do not derive  $\perp$
  - But then  $D$  cannot be correct (contradiction)

Therefore  $TI(\varepsilon_0, P)$  implies  $\forall \alpha < \varepsilon_0 P(\alpha)$ , i.e.,  $\text{Con}(PA)$ .

## 2.6 Proof of (ii): $PA \not\vdash TI(\varepsilon_0)$

**Idea:** If  $PA \vdash TI(\varepsilon_0)$ , then:

- $PA \vdash \text{Con}(PA)$  (by point (i))
- This contradicts Gödel's second incompleteness theorem

**Refinement:** Using Gödel's hierarchy of  $\Pi_1^0$  formulas, one shows that  $TI(\varepsilon_0)$  is a  $\Pi_2^0$  formula not derivable in PA.



## 2.7 Proof of (iii): $PA \vdash TI(\alpha)$ for $\alpha < \varepsilon_0$

For every  $\alpha < \varepsilon_0$ , we can encode transfinite induction up to  $\alpha$  using ordinary induction of PA, iterated a finite number of times.

**Method:** Case construction on the Cantor normal form of  $\alpha$ :

- If  $\alpha = n$  (finite):  $TI(n)$  is equivalent to ordinary induction
- If  $\alpha = \omega^{\beta_1} \cdot c_1 + \dots + \omega^{\beta_k} \cdot c_k$ : Use induction on  $\beta_1, \dots, \beta_k$  (inductive hypothesis) and iterate  $c_1, \dots, c_k$  times

The key is that each component of the representation is strictly smaller than  $\alpha$ , allowing recursive application.

## 2.8 Philosophical Consequences

Gentzen's theorem shows that:

1. **PA's consistency is provable**, but only using methods beyond PA itself
2. **The "step beyond" is minimal**:  $\varepsilon_0$  is the smallest ordinal PA cannot "see"
3. **This step is conceptually natural**: Finite vs. infinite transfinite induction
4. **Hilbert was right** that "extended finite" methods suffice, if we interpret "finite" as " $< \varepsilon_0$ "

## 2.9 Modern Formulations

Today several equivalent formulations exist:

### Type Theory

Interpretation of PA in ramified type theory up to  $\varepsilon_0$

### Set Theory

Models of PA in levels  $V_\alpha$  for  $\alpha < \varepsilon_0$

### Category Theory

Topoi of sheaves on ordinals  $< \varepsilon_0$

### Computational Theory

Termination of programs with ordinal complexity  $< \varepsilon_0$

## 2.10 The Ordinal Analysis Method

Gentzen's proof established the paradigm of **ordinal analysis**:

**Input:** A formal system  $S$

**Goal:** Find the "proof-theoretic ordinal"  $|S|$  such that:

1.  $S + TI(|S|)$  proves  $\text{Con}(S)$
2.  $S$  does not prove  $Tl(|S|)$
3.  $S$  proves  $Tl(\alpha)$  for all  $\alpha < |S|$

This method has been applied to increasingly strong systems:

|                            |                               |                           |                   |                         |                     |                  |            |
|----------------------------|-------------------------------|---------------------------|-------------------|-------------------------|---------------------|------------------|------------|
| System                     | $ S $                         | Methods Required          | ----- ----- ----- | PRA                     | $\omega^\omega$     | Combinatorial    |            |
| PA                         | $\epsilon_0$                  | Gentzen's cut-elimination | ACA <sub>0</sub>  | $\epsilon_0$            | Similar to PA       | ATR <sub>0</sub> | $\Gamma_0$ |
| $\Pi^1_1$ -CA <sub>0</sub> | $\psi_0(\epsilon_{\Omega+1})$ | Bachmann-Howard collapse  | KP                | $\psi_0(\Omega_\Omega)$ | Admissible ordinals |                  |            |

## 2.11 The Cut-Elimination Process

The heart of Gentzen's proof is the **cut-elimination procedure**:

**Cut rule:**

$$\frac{\Gamma \vdash \Delta, \varphi \quad \varphi, \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'} \text{ (Cut)}$$

**Gentzen's Hauptsatz:** Every derivation can be transformed into one without cuts.

**Connection to ordinals:**

- Each reduction step decreases an ordinal measure
- Termination guaranteed by well-foundedness up to  $\epsilon_0$
- This is where  $Tl(\epsilon_0)$  is needed!

## 2.12 Why Exactly $\epsilon_0$ ?

The ordinal  $\epsilon_0$  appears naturally because:

1. **Cantor normal form:** Every  $\alpha < \epsilon_0$  has finite representation using  $\omega, +, \cdot, ^$
2. **PA's induction:** Can handle all "finitely describable" ordinals
3. **Fixed point:**  $\epsilon_0 = \omega^{\epsilon_0}$  cannot be built from below using PA's means
4. **Closure:** Below  $\epsilon_0$  = closure under primitive recursive operations

## 2.13 Extensions and Generalizations

### Higher Ordinals for Stronger Systems

**Theorem (Takeuti, Buchholz, Rathjen):** The ordinal analysis method extends to:

- Second-order arithmetic subsystems
- Weak set theories (Kripke-Platek)
- Subsystems of ZFC
- (Partial results for full ZFC)

## Proof Complexity

**Modern question:** How does  $|S|$  relate to the length of proofs in  $S$ ?

**Result:** Systems with larger  $|S|$  can have exponentially shorter proofs for the same theorems.

## 2.14 The Predicative-Impredicative Boundary

Gentzen's work connects to the predicativity debate:

**Feferman-Schütte Theorem (1964):**

- $\Gamma_0$  is the limit of predicative mathematics
- $\varepsilon_0 < \Gamma_0$ , so PA is predicatively acceptable
- Beyond  $\Gamma_0$  requires genuinely impredicative definitions

This gives precise mathematical meaning to philosophical disputes about acceptable mathematical methods.

## 2.15 Computational Interpretation

**Modern view** (via Curry-Howard correspondence):

- Derivation  $\leftrightarrow$  Program
- Cut-elimination  $\leftrightarrow$  Program evaluation
- Ordinal  $< \varepsilon_0 \leftrightarrow$  Program termination proof

**Corollary:** Every program extracted from a PA proof terminates in time bounded by a function of ordinal  $< \varepsilon_0$ .

## 2.16 Open Problems Related to Gentzen's Work

1. **Optimal bounds:** Can we improve the complexity bounds for cut-elimination?
2. **Automated ordinal analysis:** Can we mechanize finding  $|S|$  for new systems?
3. **Non-well-founded systems:** What replaces ordinals for circular reasoning?
4. **Reverse mathematics:** For which theorems is their proof-theoretic ordinal intrinsic?

## Part 3: Beyond $\varepsilon_0$ - The Extended Veblen Hierarchy

### 3.1 Motivation: The Limitations of $\varepsilon_0$

Ordinals  $< \varepsilon_0$  are closed under:

- Ordinal addition and multiplication
- Exponentiation:  $\alpha \mapsto \omega^\alpha$
- But **NOT** under the function  $\alpha \mapsto \varepsilon_\alpha$

To study formal systems stronger than PA, we need larger ordinals with more sophisticated notation systems.

### 3.2 The Unary Veblen Function $\varphi_n$

**Recursive definition:**

$$\varphi_0(\alpha) := \omega^\alpha$$

For  $n > 0$ :  $\varphi_{n+1}(\alpha) :=$  the  $\alpha$ -th fixed point of  $\varphi_n$

where "the  $\alpha$ -th fixed point" means:

- $\varphi_{n+1}(0) = \min\{\beta : \varphi_n(\beta) = \beta\}$
- $\varphi_{n+1}(\alpha+1) = \min\{\beta > \varphi_{n+1}(\alpha) : \varphi_n(\beta) = \beta\}$
- $\varphi_{n+1}(\lambda) = \sup\{\varphi_{n+1}(\alpha) : \alpha < \lambda\}$  for limit  $\lambda$

#### Explicit Examples

**Level 0:  $\varphi_0(\alpha) = \omega^\alpha$**

- $\varphi_0(0) = 1$
- $\varphi_0(1) = \omega$
- $\varphi_0(2) = \omega^2$
- $\varphi_0(\omega) = \omega^\omega$

**Level 1:  $\varphi_1(\alpha) = \varepsilon_\alpha$**

- $\varphi_1(0) = \varepsilon_0$  (first fixed point of exponentiation)
- $\varphi_1(1) = \varepsilon_1$  (second fixed point)
- $\varphi_1(\omega) = \varepsilon_\omega$
- $\varphi_1(\varepsilon_0) = \varepsilon_{\varepsilon_0}$

**Level 2:  $\varphi_2(\alpha) =$  "critical  $\varepsilon$ -numbers"**

- $\varphi_2(0) = \zeta_0 := \min\{\beta : \varepsilon_\beta = \beta\}$
- $\varphi_2(1) = \zeta_1$
- For every  $\alpha$ ,  $\varphi_2(\alpha)$  satisfies  $\varepsilon_{\varphi_2(\alpha)} = \varphi_2(\alpha)$

**Level 3:  $\varphi_3(0) = \eta_0 := \min\{\beta : \zeta_\beta = \beta\}$**

## Pattern Recognition

Each level creates fixed points of the previous level:

- $\varphi_0$  creates powers of  $\omega$
- $\varphi_1$  creates fixed points of  $\varphi_0$  (the  $\varepsilon$ -numbers)
- $\varphi_2$  creates fixed points of  $\varphi_1$  (the  $\zeta$ -numbers)
- $\varphi_3$  creates fixed points of  $\varphi_2$  (the  $\eta$ -numbers)
- ...

## 3.3 The Binary Veblen Function $\varphi(\alpha, \beta)$

To handle arbitrary indices uniformly, Veblen introduced a two-argument function.

**Definition:**

$$\varphi(0, \beta) := \omega^\beta$$

$$\varphi(\alpha+1, \beta) := \text{the } \beta\text{-th fixed point of } \gamma \mapsto \varphi(\alpha, \gamma)$$

$$\varphi(\lambda, \beta) \text{ for limit } \lambda := \text{the } \beta\text{-th ordinal that is a simultaneous fixed point of all } \varphi(\alpha, \cdot) \text{ for } \alpha < \lambda$$

### Relation to Unary Notation

For finite  $n$ :

- $\varphi(n, \beta) = \varphi_n(\beta)$

For infinite first argument:

- $\varphi(\omega, 0) = \min\{\beta : \varphi(n, \beta) = \beta \text{ for all } n < \omega\}$

### Fundamental Properties

1. **Monotonicity:** If  $\alpha_1 \leq \alpha_2$  and  $\beta_1 \leq \beta_2$ , then  $\varphi(\alpha_1, \beta_1) \leq \varphi(\alpha_2, \beta_2)$
2. **Normality:**  $\varphi(\alpha, \cdot)$  is continuous (normal) for every  $\alpha$
3. **Fixed point generation:**  $\varphi(\alpha+1, \beta)$  is always a fixed point of  $\varphi(\alpha, \cdot)$
4. **Enumeration property:**  $\varphi(\alpha, \cdot)$  enumerates exactly the common fixed points of all  $\varphi(\gamma, \cdot)$  for  $\gamma < \alpha$
5. **Diagonal property:**  $\varphi(\alpha, 0) > \alpha$  for all  $\alpha$

## 3.4 Veblen Normal Form

**Theorem:** Every ordinal  $\alpha < \Gamma_0$  can be uniquely represented as:

$$\alpha = \varphi(\alpha_1, \beta_1) + \varphi(\alpha_2, \beta_2) + \dots + \varphi(\alpha_n, \beta_n)$$

where:

- $\varphi(\alpha_1, \beta_1) \geq \varphi(\alpha_2, \beta_2) \geq \dots \geq \varphi(\alpha_n, \beta_n)$
- Each  $\alpha_i, \beta_i < \alpha$  is itself in normal form
- Each term is **principal**:  $\varphi(\alpha_i, \beta_i)$  is not a sum of smaller ordinals

## Example Representations

**Simple examples:**

$$\varphi(1, 0) = \varepsilon_0$$

$$\varphi(1, 0) + \omega^2 + 5 = \varepsilon_0 + \omega^2 + 5$$

$$\varphi(1, 1) = \varepsilon_1$$

$$\varphi(2, 0) = \zeta_0$$

**Complex examples:**

$$\varphi(2, 3) \cdot \omega + \varphi(1, \varphi(1, 0)) + \varphi(0, \varphi(0, 7))$$

$$\varphi(\omega, 0) + \varphi(5, \varphi(2, 1)) + \omega^{\varphi(0, 3)} + 42$$

## 3.5 The Feferman-Schütte Ordinal $\Gamma_0$

**The Self-Reference Problem:** The notation  $\varphi(\alpha, \beta)$  requires  $\alpha$  and  $\beta$  to be already defined. But for large  $\alpha$ , we want to use the  $\varphi$  notation itself within the first argument.

**Question:** What is the smallest ordinal that "contains itself" in the  $\varphi$  notation?

### Definition

$$\Gamma_0 := \min\{\alpha : \varphi(\alpha, 0) = \alpha\}$$

**Equivalently:**

- $\Gamma_0 = \sup\{\varphi(0, 0), \varphi(\varphi(0, 0), 0), \varphi(\varphi(\varphi(0, 0), 0), 0), \dots\}$
- $\Gamma_0$  is the first ordinal that is a fixed point of  $\alpha \mapsto \varphi(\alpha, 0)$

### Explicit Construction

Define the fundamental sequence:

- $\gamma_0 = 0$
- $\gamma_{n+1} = \varphi(\gamma_n, 0)$
- $\Gamma_0 = \sup\{\gamma_n : n \in \mathbb{N}\}$

**Computing initial values:**

$$\gamma_0 = 0$$

$$\gamma_1 = \varphi(0, 0) = \omega^0 = 1$$

$$\gamma_2 = \varphi(1, 0) = \varepsilon_0$$

$$\gamma_3 = \varphi(\varepsilon_0, 0) = \zeta_0 \text{ (first fixed point of } \alpha \mapsto \varepsilon_\alpha \text{)}$$

$$\gamma_4 = \varphi(\zeta_0, 0)$$

$$\gamma_5 = \varphi(\varphi(\zeta_0, 0), 0)$$

...

Each  $\gamma_n$  is vastly larger than the previous, yet  $\Gamma_0$  transcends them all.

**Mathematical Significance of  $\Gamma_0$** **Theorem (Feferman-Schütte, 1964):**

1.  $\Gamma_0$  is the exact limit of predicative mathematics
2.  $|\text{ATR}_0| = \Gamma_0$
3. Beyond  $\Gamma_0$  requires genuinely impredicative definitions

**Philosophical meaning:**

- Below  $\Gamma_0$ : "generalized finitism"
- At  $\Gamma_0$ : predicative ceiling
- Above  $\Gamma_0$ : essential impredicativity

**3.6 The Extended Veblen Hierarchy**

To go beyond  $\Gamma_0$ , Veblen introduced functions of arbitrary finite arity.

**n-ary Veblen Functions**

$$\varphi(\alpha_1, \alpha_2, \dots, \alpha_n, \beta):$$

$$\text{Base case: } \varphi(\beta) := \omega^\beta$$

**Recursive case:**  $\varphi(\alpha_1, \dots, \alpha_n, \beta) :=$  the  $\beta$ -th common fixed point of all  $\varphi(\alpha_1, \dots, \alpha_{n-1}, \cdot)$  where  $\alpha_n < \alpha_n$

**Notable Ordinals in the Extended Hierarchy****Using ternary Veblen:**

- $\Gamma_0 = \varphi(1, 0, 0)$
- $\Gamma_1 = \varphi(1, 0, 1)$
- $\Gamma_\omega = \varphi(1, 0, \omega)$
- $\Gamma_-(\Gamma_0) = \varphi(1, 0, \Gamma_0)$

**Next level:**

- $\varphi(1, 1, 0)$  = first fixed point of  $\alpha \mapsto \varphi(1, 0, \alpha)$
- $\varphi(1, 2, 0)$  = first fixed point of  $\alpha \mapsto \varphi(1, 1, \alpha)$
- $\varphi(1, \omega, 0)$  = first common fixed point of all  $\varphi(1, n, \cdot)$

**Even higher:**

- $\varphi(2, 0, 0)$  = first fixed point of  $\alpha \mapsto \varphi(1, \alpha, 0)$
- $\varphi(\omega, 0, 0)$  = first common fixed point of all  $\varphi(n, \cdot, \cdot)$

**The Small Veblen Ordinal (SVO)**

**Definition:** SVO = supremum of all ordinals expressible using finitely many arguments to  $\varphi$

$$\text{SVO} = \sup\{\varphi(\alpha_1, \dots, \alpha_n, \beta) : n < \omega, \text{ all arguments expressible}\}$$

**Properties:**

- SVO cannot be expressed using Veblen functions alone
- SVO marks the limit of "transparent" ordinal notations
- Beyond SVO requires fundamentally new techniques

**3.7 Ordinals and Formal Systems: The Correspondence****Fundamental table:**

| Ordinal                          | System                | Characteristic Principle            |
|----------------------------------|-----------------------|-------------------------------------|
| $\omega^\omega$                  | PRA                   | Primitive recursive induction       |
| $\varepsilon_0$                  | PA                    | First-order induction               |
| $\varphi(\omega, 0)$             | $\text{ACA}_0$        | Arithmetical comprehension          |
| $\Gamma_0$                       | $\text{ATR}_0$        | Arithmetical transfinite recursion  |
| $\psi_0(\varepsilon_{\Omega+1})$ | $\Pi^1_1\text{-CA}_0$ | $\Pi^1_1$ -comprehension            |
| $\psi(\Omega_\omega)$            | $\text{ID}_1$         | Inductive definitions iterated once |
| $\psi(\Omega_\Omega)$            | KP                    | Kripke-Platek set theory            |

**Pattern:** Stronger system  $\Leftrightarrow$  Larger ordinal  $\Leftrightarrow$  More complex normalization

**3.8 The Bachmann-Howard Collapsing Function  $\psi$**



To handle ordinals beyond SVO, we need a radically different approach.

## The Collapsing Idea

**Problem:** Veblen functions eventually run out of "naming power"

**Solution:** Use **uncountable ordinals** as temporary scaffolding, then "collapse" them back to countable ordinals

## Uncountable Ordinal Markers

- $\Omega$  = first uncountable ordinal ( $\omega_1$ )
- $\Omega_2$  = second uncountable ordinal ( $\omega_2$ )
- $\Omega_\omega$  =  $\omega$ -th uncountable ordinal
- ...

## The $\psi$ Function

**Bachmann-Howard collapsing function**  $\psi_\Omega(\alpha)$ :

The function  $\psi_\Omega(\alpha)$  denotes the  $\alpha$ -th ordinal that is **constructible** using only:

1. 0, +
2. Veblen  $\varphi$  functions
3. Ordinals  $< \Omega$
4. Previous results of  $\psi_\Omega$

**Key insight:** We use  $\Omega$  in the "construction phase" but the resulting ordinal is countable ( $< \Omega$ ).

## Precise Definition

Define the constructible closure:

$$C_0(\alpha, \beta) = \beta \cup \{0, \Omega\}$$

$$C_{n+1}(\alpha, \beta) = C_n(\alpha, \beta) \cup \{\gamma + \delta, \varphi(\gamma, \delta), \psi_\xi(\eta) : \gamma, \delta, \xi, \eta \in C_n(\alpha, \beta), \xi < \alpha\}$$

$$C(\alpha, \beta) = \bigcup_{n < \omega} C_n(\alpha, \beta)$$

$$\psi_\alpha(\beta) = \min\{\gamma < \Omega : \beta \in C(\alpha, \gamma) \wedge \gamma \notin C(\alpha, \gamma)\}$$

## Interpretation

- $C(\alpha, \beta)$  is the closure of  $\beta$  under operations
- $\psi_\alpha(\beta)$  is the smallest ordinal  $\gamma$  such that:
  - $\beta$  can be built using things  $< \gamma$
  - But  $\gamma$  itself cannot be built (collapsing point)

## Computing Examples

$\psi_0(0) = \Omega$  (nothing smaller can construct 0 with  $\Omega$ )

$\psi_0(1) = \varepsilon_{\Omega+1}$

$\psi_{\Omega}(0) \approx \Gamma_0$

$\psi_{\Omega}(\varepsilon_{\Omega+1}) = \text{ordinal of } \Pi^1_1\text{-CA}_0$

## 3.9 General Ordinal Collapsing Technique

**Universal pattern** for denoting large ordinals:

1. **Expand:** Work in universe with large (uncountable) ordinals
2. **Build:** Construct expressions using these large ordinals
3. **Collapse:** Remove references to uncountable ordinals, obtaining countable result

### Examples Beyond Bachmann-Howard

**Using multiple uncountables:**

- $\psi_{\{\Omega_2\}}(\alpha)$  uses both  $\Omega$  and  $\Omega_2$
- $\psi_{\{\Omega_\omega\}}(\alpha)$  uses the sequence  $\Omega, \Omega_2, \Omega_3, \dots$

**For large cardinal strength:**

- Use I (first inaccessible) as marker
- Use M (first Mahlo) as marker
- Use K (first weakly compact) as marker

## 3.10 Connection with Proof Theory

These large ordinals measure the complexity of proofs in increasingly strong systems.

**The fundamental pattern:**

- Stronger logical system
- $\Rightarrow$  Larger proof-theoretic ordinal
- $\Rightarrow$  More complex cut-elimination
- $\Rightarrow$  More powerful normalization techniques

**Hilbert's program continues:** Even after Gödel's incompleteness theorems, ordinal analysis extends Hilbert's vision by:

- Measuring consistency strength precisely
- Finding constructive content in classical proofs
- Calibrating logical systems on absolute scale

## 3.11 Examples of Large Ordinal Computations

### System Ordinals

$\Pi^1_1\text{-CA}_0$ :  $\psi_0(\varepsilon_{\Omega+1})$

- Uses  $\Omega$  as marker
- Collapses back to countable
- Far beyond  $\Gamma_0$

$\Delta^1_2\text{-CA}_0 + \text{BI}$ :  $\psi_0(\varepsilon_{\Omega_\omega+1})$

- Uses sequence of  $\Omega$ s
- Even more complex

$\Pi^1_2\text{-CA}_0$ :  $\psi_0(\varepsilon_{M+1})$

- Uses Mahlo cardinal  $M$
- Reflects large cardinal strength

## 3.12 Philosophical Significance

The ordinal hierarchy reveals fundamental structure:

### Natural Plateaus:

- $\varepsilon_0$ : Extended finitism
- $\Gamma_0$ : Predicativity limit
- $\psi_0(\varepsilon_{\Omega+1})$ : Impredicativity threshold
- $\omega_1^{\text{CK}}$ : Recursiveness limit

**Not Arbitrary:** These levels reflect deep structural properties of mathematical reasoning, not human convention.

## 3.13 Open Problems

1. **Optimal notations:** Are there better notation systems beyond Bachmann-Howard?
2. **Computational complexity:** Precise relationship between ordinal  $\alpha$  and complexity of functions with termination ordinal  $\alpha$ ?
3. **Natural intermediates:** Are there "natural" formal systems with ordinals between major landmarks?
4. **Large cardinal reach:** How far can ordinal analysis extend into large cardinal territory?

5. **Alternative foundations:** Do ordinals remain meaningful in constructive or category-theoretic foundations?
6. **ZFC ordinal:** What is  $|ZFC|$ ? Only bounds are known.

## Part 4: Reverse Mathematics - The Complete Hierarchy

### 4.1 Philosophy and Motivation

**Fundamental question:** What axioms are **necessary** to prove a given mathematical theorem?

**Reverse Mathematics** (inverse mathematics): Instead of deriving theorems from axioms, find the **minimal axioms** necessary for a theorem.

#### The Method

1. Fix a weak base system (typically  $RCA_0$ )
2. For a theorem  $T$ , prove:  $\text{Base} + T \rightarrow \text{Axiom } A$
3. If also  $\text{Base} + A \rightarrow T$ , then  $T$  and  $A$  are **equivalent**

**Surprising phenomenon** (Friedman-Simpson, 1970s): Most ordinary mathematical theorems are equivalent to one of five systems ("The Big Five").

### 4.2 The Base System: $RCA_0$

#### Recursive Comprehension Axioms

#### Language $L_2$ (Second-Order Arithmetic)

- **Number variables:**  $n, m, i, j, \dots$
- **Set variables:**  $X, Y, Z, \dots$
- **Symbols:**  $0, S, +, \times, <, \in$

#### Axioms of $RCA_0$

##### 1. Basic arithmetic axioms: $PA^-$ (PA without induction)

- Equations for  $+$ ,  $\times$
- Linear ordering of  $<$
- Discreteness:  $n < n+1$

##### 2. $\Sigma^0_1$ -induction: For every $\Sigma^0_1$ formula $\varphi(n)$ :

$$[\varphi(0) \wedge \forall n(\varphi(n) \rightarrow \varphi(n+1))] \rightarrow \forall n \varphi(n)$$

Limited to formulas with bounded existential quantifiers

**3. Recursive comprehension:** For every  $\Delta^0_1$  formula  $\varphi(n)$ :

$$\forall n [\varphi(n) \leftrightarrow \psi(n)] \rightarrow \exists X \forall n [n \in X \leftrightarrow \varphi(n)]$$

where  $\varphi$  is  $\Sigma^0_1$  and  $\psi$  is  $\Pi^0_1$

**Interpretation:**  $\text{RCA}_0$  formalizes "algorithmic" mathematics—sets constructible by recursive algorithms.

**Minimal model:**  $(\mathbb{N}, \text{REC})$  where  $\text{REC}$  = recursive sets of natural numbers

**Proof-theoretic ordinal:**  $|\text{RCA}_0| = \omega^\omega$

## 4.3 $\text{WKL}_0$ : Weak König's Lemma

**Axioms:**  $\text{RCA}_0 + \text{WKL}$

### Weak König's Lemma

**Statement:** Every infinite binary tree has an infinite branch.

**Formally:** If  $T \subseteq \{0,1\}^{<\mathbb{N}}$  is a tree (closed under initial segments) and  $\forall n \exists \sigma \in T (|\sigma| = n)$ , then  $\exists f: \mathbb{N} \rightarrow \{0,1\}$  such that  $\forall n (f \upharpoonright n \in T)$ .

### Theorems Equivalent to $\text{WKL}_0$ (over $\text{RCA}_0$ )

**Analysis and Topology:**

1. **Heine-Borel theorem:** Every open covering of  $[0,1]$  has a finite subcovering
2. **Brouwer's theorem:** Every continuous function  $f: [0,1] \rightarrow \mathbb{R}$  is bounded
3. **Intermediate value theorem:** If  $f: [0,1] \rightarrow \mathbb{R}$  continuous with  $f(0) < 0 < f(1)$ , then  $\exists c: f(c) = 0$
4. **Fundamental theorem of algebra:** Every non-constant polynomial in  $\mathbb{C}$  has a root

**Logic:** 5. **Gödel's completeness theorem:** Every consistent theory has a countable model

6. **Compactness theorem:** A finitely consistent theory is consistent

**Set Theory:** 7.  **$\Pi^0_1$  separation:** If  $A, B$  are disjoint  $\Pi^0_1$  sets,  $\exists C$  recursive separating  $A$  and  $B$

**Minimal model:**  $(\mathbb{N}, \text{REC} \cup \text{Low sets})$  where  $\text{Low} = \{X : X' \leq_T 0'\}$

**Proof-theoretic ordinal:**  $|\text{WKL}_0| = \omega^\omega$  (same as  $\text{RCA}_0$ !)

**Computational characterization:**  $WKL_0$  doesn't increase computational complexity but allows compactness arguments.

## 4.4 $ACA_0$ : Arithmetical Comprehension

**Axioms:**  $RCA_0 + ACA$

### Arithmetical Comprehension

For every arithmetical formula  $\varphi(n)$ :

$$\exists X \forall n [n \in X \leftrightarrow \varphi(n)]$$

**Interpretation:** Every arithmetical property (even with quantifiers over numbers) defines a set.

### Theorems Equivalent to $ACA_0$

**Analysis and Topology:**

1. **Bolzano-Weierstrass theorem:** Every bounded sequence has a convergent subsequence
2. **Bolzano's ascending theorem:** Every bounded monotone sequence converges
3. **General Heine-Borel:** Every closed and bounded set is compact

**Algebra:** 4. **König's theorem for graphs:** An infinite locally finite graph has an infinite matching 5. **Decomposition of abelian groups:** Basic theorems on countable abelian groups 6. **Existence of maximal ideals:** In countable commutative rings

**Combinatorics:** 7. **Infinite Ramsey theorem:**  $\forall c: [\mathbb{N}]^2 \rightarrow \{0,1\}, \exists H$  infinite homogeneous 8. **Dickson's lemma:**  $(\mathbb{N}^n, \leq)$  is well-quasi-ordered

**Logic:** 9. **Consistency implies  $\omega$ -consistency:** For recursively axiomatizable theories 10. **Turing-Feferman theorem:** Turing jump  $X'$  exists for every  $X$

**Minimal model:**  $(\mathbb{N}, ARITH)$  where  $ARITH$  = complete arithmetical hierarchy

**Proof-theoretic ordinal:**  $|ACA_0| = \varepsilon_0$

**Characterization:**  $ACA_0$  is exactly PA in second-order formulation.

## 4.5 $ATR_0$ : Arithmetical Transfinite Recursion

**Axioms:**  $RCA_0 + ATR$

### Arithmetical Transfinite Recursion

For every arithmetical formula  $\varphi$ :

If  $<$  is a well-ordering, then  $\exists F: \forall \alpha [F(\alpha) = \{n : \varphi(\alpha, F \restriction \alpha, n)\}]$

where  $F \restriction \alpha = \{(\beta, n) : \beta < \alpha \wedge n \in F(\beta)\}$

**Interpretation:** Allows definitions by transfinite recursion along arithmetical well-orderings.

## Theorems Equivalent to $\text{ATR}_0$

**Ordinal Theory:**

1. **Comparability of well-orderings:** For all well-orderings  $<_1, <_2$ , there exists an isomorphism or embedding
2. **Existence of rank functions:** Every well-ordering has an associated ordinal
3. **Representability:** Every countable ordinal is representable

**Topology:** 4. **Cantor-Bendixson theorem:** Every Polish space has a perfect derivation until it becomes perfect or empty 5. **Perfect set theorem for closed sets:** Every uncountable closed set contains a perfect subset

**Descriptive Set Theory:** 6.  **$\Sigma^1_1$ -separation:** If  $A, B$  are disjoint  $\Sigma^1_1$  sets,  $\exists C$  arithmetical separating them 7. **Reduction principle for  $\Sigma^1_1$ :** A collection of  $\Sigma^1_1$  sets has a  $\Sigma^1_1$  reduction

**Game Theory:** 8. **Determinacy of open games:** Gale-Stewart games with open winning set are determined

**Algebra:** 9. **Ulm's theorem (base form):** Countable abelian groups have well-defined Ulm invariants 10. **Embedding theorems** for countable ordered structures

**Minimal model:**  $(\mathbb{N}, \text{HYP})$  where  $\text{HYP}$  = hyperarithmetical hierarchy =  $\Delta^1_1$  sets

**Proof-theoretic ordinal:**  $|\text{ATR}_0| = \Gamma_0$

**Characterization:**  $\text{ATR}_0$  is the maximal predicative system + transfinite recursion.

## 4.6 $\Pi^1_1\text{-CA}_0$ : $\Pi^1_1$ Comprehension

**Axioms:**  $\text{RCA}_0 + \Pi^1_1\text{-CA}$

### $\Pi^1_1$ -Comprehension

For every formula of the form  $\forall X \varphi(n, X)$  with  $\varphi$  arithmetical:

$\exists Y \forall n [n \in Y \leftrightarrow \forall X \varphi(n, X)]$

**Interpretation:** Definitions that quantify over **all** sets of numbers—genuinely impredicative.

## Theorems Equivalent to $\Pi^1_1\text{-CA}_0$

### Topology and Analysis:

1. Every non-empty analytic set contains a  $\Pi^1_1$  singleton
2. Principle of dependent choice for  $\Sigma^1_1$
3. Existence of ultrafilters (countable versions)

**Descriptive Set Theory:** 4. **Perfect set property for  $\Sigma^1_1$ :** Every uncountable  $\Sigma^1_1$  set contains a perfect subset 5. **Uniformization for  $\Sigma^1_1$ :** Every  $\Sigma^1_1$  relation has a  $\Sigma^1_1$  function witnessing it 6. **Prewellordering property for  $\Sigma^1_1$**

**Determinacy:** 7. **Determinacy for  $\Sigma^0_3$  games:** Extension beyond open games 8. **Quasi-determinacy for Borel sets**

**Set Theory:** 9. **Comparison theorem for  $\Sigma^1_1$  cardinality** 10. **Silver's theorem (weak forms):** For  $\Sigma^1_1$  equivalence

**Minimal model:**  $(\mathbb{N}, \Pi^1_1) =$  all  $\Pi^1_1$ -definable sets

**Proof-theoretic ordinal:**  $|\Pi^1_1\text{-CA}_0| = \psi_0(\varepsilon_{\Omega+1})$  (requires Bachmann-Howard notation)

**Characterization:** First genuinely impredicative system, beyond the  $\Gamma_0$  barrier.

## 4.7 The Big Five: Summary Table

| System                | Principle                  | Ordinal                          | Minimal Model  | Characterization        |
|-----------------------|----------------------------|----------------------------------|----------------|-------------------------|
| $\text{RCA}_0$        | Basic recursion            | $\omega^\omega$                  | REC            | Algorithmic mathematics |
| $\text{WKL}_0$        | Weak König                 | $\omega^\omega$                  | $\text{Low}_1$ | Finite compactness      |
| $\text{ACA}_0$        | Arithmetical comprehension | $\varepsilon_0$                  | ARITH          | PA in second order      |
| $\text{ATR}_0$        | Transfinite recursion      | $\Gamma_0$                       | HYP            | Maximal predicative     |
| $\Pi^1_1\text{-CA}_0$ | $\Pi^1_1$ -comprehension   | $\psi_0(\varepsilon_{\Omega+1})$ | $\Pi^1_1$      | Impredicative           |

## 4.8 Intermediate Systems

**Between  $\text{RCA}_0$  and  $\text{WKL}_0$ :**

- $\text{RCA}_0^+$ :  $\text{RCA}_0$  + " $\Sigma^0_1$  induction with set parameters"



- DNR:  $\text{RCA}_0$  + "diagonally non-recursive functions exist"

#### Between $\text{WKL}_0$ and $\text{ACA}_0$ :

- $\text{WKL}_0^+$ :  $\text{WKL}_0$  + weak forms of comprehension
- Very few natural theorems here!

#### Between $\text{ACA}_0$ and $\text{ATR}_0$ :

- $\text{ACA}_0^+$ :  $\text{ACA}_0$  +  $\Sigma^1_1$ -induction
- Extremely few natural theorems!

#### Between $\text{ATR}_0$ and $\Pi^1_1\text{-CA}_0$ :

- $\Delta^1_1\text{-CA}_0$ : Comprehension limited to  $\Delta^1_1$
- Some theorems on automata and formal languages

#### Beyond $\Pi^1_1\text{-CA}_0$ :

- $\Pi^1_2\text{-CA}_0$ : Ordinal  $\psi_0(\varepsilon_{\{M+1\}})$  with  $M$  = first Mahlo
- $\Delta^1_2\text{-CA}_0$  + BI: Bar induction, even larger ordinal
- $Z_2$ : Full second-order arithmetic

## 4.9 The Big Five Phenomenon

**Empirical observation** (Friedman, Simpson, Seetapun):

Almost every ordinary mathematical theorem is equivalent (over  $\text{RCA}_0$ ) to one of the Big Five.

### "Homeless" Theorems (rare exceptions)

1. Some variants of Ramsey theory ( $\text{RT}^2_2$  doesn't equal any!)
2. Theorems on relativized computability
3. Some artificial combinatorial propositions

### Philosophical Explanation (Simpson)

The Big Five correspond to "natural conceptions" of sets of numbers:

- **$\text{RCA}_0$** : Algorithmically definable sets
- **$\text{WKL}_0$** : Sets from finite processes + compactness
- **$\text{ACA}_0$** : Arithmetically definable sets
- **$\text{ATR}_0$** : Sets constructible by predicative transfinite induction
- **$\Pi^1_1\text{-CA}_0$** : Impredicatively definable sets

## 4.10 Calibration: Classical Theorems

### Theorems Equivalent to $\text{RCA}_0$

- Existence of convergent monotone sequences
- Calculation with rational numbers
- Basic arithmetic theorems
- Existence of minima in finite subsets

### Theorems Equivalent to $\text{WKL}_0$

- Intermediate value theorem
- Heine-Borel theorem
- Fundamental theorem of algebra
- Compactness theorem in logic
- Every separable space has countable base

### Theorems Equivalent to $\text{ACA}_0$

- Bolzano-Weierstrass theorem
- Every sequence has supremum
- Existence of Lebesgue measure
- König's lemma
- Infinite Ramsey theorem

### Theorems Equivalent to $\text{ATR}_0$

- Cantor-Bendixson theorem
- Determinacy of open games
- Perfect set theorem for closed sets
- Comparability of countable ordinals
- Ulm's theorem (base version)

### Theorems Equivalent to $\Pi^1_1\text{-CA}_0$

- Perfect set theorem for analytic sets
- Uniformization principle
- Determinacy for complex games
- Silver's dichotomy (weak versions)

## 4.11 Conservativity and Relative Strength

#### Conservativity theorems:

1.  **$\text{WKL}_0$  over  $\text{RCA}_0$** :  $\text{WKL}_0$  is  $\Pi^0_2$ -conservative over  $\text{RCA}_0$ 
  - Every  $\Pi^0_2$  statement provable in  $\text{WKL}_0$  is already provable in  $\text{RCA}_0$
2.  **$\text{WKL}_0$  and  $\text{ACA}_0$** : Incomparable

- Neither implies the other over  $\text{RCA}_0$
- 3.  **$\text{ACA}_0$  over  $\text{RCA}_0$** : Not conservative, but  $\Pi^1_1$ -conservative
- 4.  **$\text{ATR}_0$  over  $\text{ACA}_0$** : Genuinely stronger

#### Proof-theoretic separation:

- $|\text{RCA}_0| = |\text{WKL}_0| = \omega^\omega$  (same proof-theoretic strength!)
- $|\text{ACA}_0| = \varepsilon_0$  (huge jump)
- $|\text{ATR}_0| = \Gamma_0$  (another jump)
- $|\Pi^1_1\text{-CA}_0| = \psi_0(\varepsilon_{\Omega+1})$  (explosion)

## 4.12 Reverse Mathematics for Other Areas

### Abstract Algebra

- **Group theory**: Mostly in  $\text{ACA}_0$  or  $\text{ATR}_0$
- **Ring theory**: Base in  $\text{WKL}_0$  (compactness), developments in  $\text{ACA}_0$
- **Field theory**: Much in  $\text{WKL}_0$  (via compactness)

### Functional Analysis

- **Hilbert spaces**: Base theorems in  $\text{WKL}_0$
- **Fixed point theorems**: Brouwer in  $\text{WKL}_0$ , Schauder in  $\text{ACA}_0$
- **Spectral theory**:  $\text{ATR}_0$  for general results

### Topology

- **General topology**: Much in  $\text{WKL}_0$
- **Dimension theory**:  $\text{ATR}_0$
- **Complete metric spaces**: Cantor-Bendixson in  $\text{ATR}_0$

### Measure Theory

- **Lebesgue measure**: Existence in  $\text{ACA}_0$
- **Fubini's theorem**:  $\text{ACA}_0$
- **Radon-Nikodym theorem**:  $\Pi^1_1\text{-CA}_0$

### Ergodic Theory

- **Birkhoff ergodic theorem**:  $\text{ATR}_0$
- **Recurrence theorems**:  $\text{WKL}_0$  -  $\text{ACA}_0$

## 4.13 Computational Applications

**Program extraction:** From proof in system  $S$ , extract algorithm with complexity related to  $|S|$ .

**Examples:**

- Proof in  $WKL_0$ : primitive recursive algorithm
- Proof in  $ACA_0$ : complexity  $\leq f_{\epsilon_0}$
- Proof in  $ATR_0$ : complexity  $\leq f_{\Gamma_0}$

**Automated theorem proving:** Knowing a theorem is in  $WKL_0$  vs  $ATR_0$  guides search strategy.

**Formal verification:** Systems like Coq calibrated to extract programs from different hierarchy levels.

## 4.14 Reverse Mathematics and Philosophy of Mathematics

**Philosophical question:** What are the real numbers "really"?

**Answers via Reverse Mathematics:**

- **$RCA_0$ :** Reals as recursive Cauchy sequences (strict constructivism)
- **$WKL_0$ :** Reals + compactness (extended finitism)
- **$ACA_0$ :** Arithmetically definable reals (weak predicativism)
- **$ATR_0$ :** Reals constructible by transfinite induction (full predicativism)
- **$\Pi^1_1\text{-}CA_0$ :** Impredicatively definable reals (moderate platonism)
- **$Z_2$ :** Classical reals (full platonism)

**Consequence:** "The reals" is not a univocal concept—depends on which axioms you accept!

## 4.15 Open Problems in Reverse Mathematics

1. **Gap problem:** Why so few theorems between the Big Five? Hidden structure?
2.  **$RT^2_2$  problem:** Where exactly is  $RT^2_2$  (Ramsey theorem for pairs)? Above  $RCA_0$ , below  $ACA_0$ , but equivalent to what?
3. **Natural intermediate systems:** Are there "natural" systems between  $WKL_0$  and  $ACA_0$ ?
4. **Higher-order RM:** What happens for  $Z_3, Z_4, \dots$ ? Are there "Big Five" at each level?
5. **Constructive RM:** How does the hierarchy change in intuitionistic context?

6. **Effective RM:** Precise connection with computable analysis?
7. **Category-theoretic RM:** Formulation in terms of topos theory?

## Part 5: Cut-Elimination Techniques - Complete Technical Analysis

### 5.1 The Sequent Calculus LK: Formal Definition

**Sequent:**  $\Gamma \vdash \Delta$  where  $\Gamma, \Delta$  are multisets of formulas

- **Reading:** "From hypotheses  $\Gamma$  follows at least one of the conclusions  $\Delta$ "
- **Semantics:**  $(\wedge \Gamma) \rightarrow (\vee \Delta)$

#### Structural Rules

**Axiom:**

$$\frac{}{\varphi \vdash \varphi} \text{ (Ax)}$$

**Weakening:**

$$\frac{\Gamma \vdash \Delta}{\varphi, \Gamma \vdash \Delta} \text{ (W-L)} \quad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, \varphi} \text{ (W-R)}$$

**Contraction:**

$$\frac{\varphi, \varphi, \Gamma \vdash \Delta}{\varphi, \Gamma \vdash \Delta} \text{ (C-L)} \quad \frac{\Gamma \vdash \Delta, \varphi, \varphi}{\Gamma \vdash \Delta, \varphi} \text{ (C-R)}$$

**Exchange:** (implicit in multisets)

#### Logical Rules

**Conjunction:**

$$\frac{\Gamma \vdash \Delta, \varphi \quad \Gamma \vdash \Delta, \psi}{\Gamma \vdash \Delta, \varphi \wedge \psi} \text{ (}\wedge\text{-R)} \quad \frac{\varphi, \Gamma \vdash \Delta}{\varphi \wedge \psi, \Gamma \vdash \Delta} \text{ (}\wedge\text{-L}_i\text{)}$$

$$\frac{\psi, \Gamma \vdash \Delta}{\varphi \wedge \psi, \Gamma \vdash \Delta} (\wedge\text{-L}_2)$$

**Disjunction:**

$$\frac{\varphi, \Gamma \vdash \Delta \quad \psi, \Gamma \vdash \Delta \quad \Gamma \vdash \Delta, \varphi}{\varphi \vee \psi, \Gamma \vdash \Delta \quad \Gamma \vdash \Delta, \varphi \vee \psi} (\vee\text{-L}) \quad \text{---} (\vee\text{-R}_1)$$

$$\frac{\Gamma \vdash \Delta, \psi}{\Gamma \vdash \Delta, \varphi \vee \psi} (\vee\text{-R}_2)$$

**Implication:**

$$\frac{\Gamma \vdash \Delta, \varphi \quad \psi, \Gamma' \vdash \Delta' \quad \varphi, \Gamma \vdash \Delta, \psi}{\varphi \rightarrow \psi, \Gamma, \Gamma' \vdash \Delta, \Delta' \quad \Gamma \vdash \Delta, \varphi \rightarrow \psi} (\rightarrow\text{-L}) \quad \text{---} (\rightarrow\text{-R})$$

**Quantifiers:**

$$\frac{\varphi(t), \Gamma \vdash \Delta \quad \Gamma \vdash \Delta, \varphi(y)}{\forall x \varphi(x), \Gamma \vdash \Delta \quad \Gamma \vdash \Delta, \forall x \varphi(x)} (\forall\text{-L}) \quad \text{---} (\forall\text{-R}) [y \text{ fresh}]$$

$$\frac{\varphi(y), \Gamma \vdash \Delta [y \text{ fresh}] \quad \Gamma \vdash \Delta, \varphi(t)}{\exists x \varphi(x), \Gamma \vdash \Delta \quad \Gamma \vdash \Delta, \exists x \varphi(x)} (\exists\text{-L}) \quad \text{---} (\exists\text{-R})$$

**Cut Rule:**

$$\frac{\Gamma \vdash \Delta, \varphi \quad \varphi, \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'} (\text{Cut})$$

## 5.2 Complexity Measures for Derivations

For a derivation D:

### 1. Height h(D)

$$h(\text{axiom}) = 0$$

$$h(\text{unary rule } D'/D) = h(D') + 1$$

$$h(\text{binary rule } D_1, D_2/D) = \max(h(D_1), h(D_2)) + 1$$

## 2. Degree $g(D)$

$g(D)$  = maximum logical complexity of formulas in  $D$

Complexity of  $\varphi$ :

- $|\perp| = |\top| = 0$
- $|P(t)| = 1$  for atomic
- $|\varphi \circ \psi| = \max(|\varphi|, |\psi|) + 1$  for  $\circ \in \{\wedge, \vee, \rightarrow\}$
- $|Qx \varphi| = |\varphi| + 1$  for  $Q \in \{\forall, \exists\}$

## 3. Cut-rank $cr(D)$

$cr(D)$  = maximum complexity of cut formulas in  $D$

$cr(D \text{ without cuts}) = 0$

## 4. Multiplicity $m(D)$

$m(D)$  = total number of Cut rule applications in  $D$

## 5. Cut-height $ch(D)$

$ch(\text{cut on } \varphi \text{ at height } k) = k$

$ch(D)$  = maximum cut-height in  $D$

# 5.3 Elimination Algorithm: Detailed Procedure

**Input:** Derivation  $D$  of  $\Gamma \vdash \Delta$  with  $cr(D) = r > 0$

**Output:** Derivation  $D'$  of  $\Gamma \vdash \Delta$  with  $cr(D') < r$

## Phase 1: Isolate Maximum Rank Cut

Find the highest cut with formula of complexity  $r$ :

$$\begin{array}{c}
 \begin{array}{cc}
 D_1 & D_2 \\
 \hline
 \Gamma \vdash \Delta, \varphi & \varphi, \Gamma' \vdash \Delta'
 \end{array} \\
 \hline
 \Gamma, \Gamma' \vdash \Delta, \Delta' \quad \text{(Cut on } \varphi, |\varphi| = r)
 \end{array}$$

## Phase 2: Case Analysis on $\varphi$

**Case A:  $\varphi$  is Atomic****Subcase A1:**  $\varphi$  is axiomatic in  $D_1$  (i.e.,  $\varphi \in \Gamma$ ) $D_1$  contains  $\varphi$  in hypotheses  $\rightarrow \varphi \vdash \varphi$  by axiom

Eliminate cut by replacing with direct axiom.

**Complexity:**  $O(1)$  - immediate elimination**Case B:  $\varphi = \psi \wedge \chi$**  $D_1$  must end with  $(\wedge\text{-R})$ :

$$\begin{array}{c}
\begin{array}{c} D_1': \\ \Gamma \vdash \Delta, \psi \end{array} \quad \begin{array}{c} D_1'': \\ \Gamma \vdash \Delta, \chi \end{array} \\
\hline
\Gamma \vdash \Delta, \psi \wedge \chi
\end{array}
\quad (\wedge\text{-R})$$

 $D_2$  must use  $(\wedge\text{-L}_1)$  or  $(\wedge\text{-L}_2)$ :**If  $D_2$  uses  $(\wedge\text{-L}_1)$ :**

$$\begin{array}{c}
\begin{array}{c} D_2': \\ \psi, \Gamma' \vdash \Delta' \end{array} \\
\hline
\psi \wedge \chi, \Gamma' \vdash \Delta'
\end{array}
\quad (\wedge\text{-L}_1)$$

**Reduction:**

$$\begin{array}{c}
\begin{array}{c} D_1' \\ \Gamma \vdash \Delta, \psi \end{array} \quad \begin{array}{c} D_2' \\ \psi, \Gamma' \vdash \Delta' \end{array} \\
\hline
\Gamma, \Gamma' \vdash \Delta, \Delta'
\end{array}
\quad (\text{Cut on } \psi, |\psi| < r)$$

**If  $D_2$  uses  $(\wedge\text{-L}_2)$ :** Symmetric with  $\chi$ **Complexity:** One new cut of rank  $< r$ **Case C:  $\varphi = \psi \vee \chi$**  $D_1$  must use  $(\vee\text{-R}_1)$  or  $(\vee\text{-R}_2)$ :

$$\begin{array}{c}
\begin{array}{c} D_1': \\ \Gamma \vdash \Delta, \psi \end{array} \\
\hline
\Gamma \vdash \Delta, \psi \vee \chi
\end{array}
\quad (\vee\text{-R}_1)$$



$D_2$  must end with ( $\vee$ -L):

$$\frac{\begin{array}{c} D_2': \\ \psi, \Gamma' \vdash \Delta' \end{array} \quad \begin{array}{c} D_2'': \\ \chi, \Gamma' \vdash \Delta' \end{array}}{\psi \vee \chi, \Gamma' \vdash \Delta'} \quad (\vee\text{-L})$$

**Reduction:**

$$\frac{\begin{array}{c} D_1': \\ \Gamma \vdash \Delta, \psi \end{array} \quad \begin{array}{c} D_2': \\ \psi, \Gamma' \vdash \Delta' \end{array}}{\Gamma, \Gamma' \vdash \Delta, \Delta'} \quad (\text{Cut on } \psi)$$

**If  $D_1$  uses ( $\vee$ -R<sub>2</sub>):** Use  $D_2''$  instead

**Complexity:** One new cut of rank  $< r$

**Case D:  $\varphi = \psi \rightarrow \chi$**

$D_1$  ends with ( $\rightarrow$ -R):

$$\frac{\begin{array}{c} D_1': \\ \psi, \Gamma \vdash \Delta, \chi \end{array}}{\Gamma \vdash \Delta, \psi \rightarrow \chi} \quad (\rightarrow\text{-R})$$

$D_2$  ends with ( $\rightarrow$ -L):

$$\frac{\begin{array}{c} D_2': \\ \Gamma' \vdash \Delta', \psi \end{array} \quad \begin{array}{c} D_2'': \\ \chi, \Gamma'' \vdash \Delta'' \end{array}}{\psi \rightarrow \chi, \Gamma', \Gamma'' \vdash \Delta', \Delta''} \quad (\rightarrow\text{-L})$$

**Reduction** (more complex):

$$\frac{\begin{array}{c} D_1' \\ \psi, \Gamma \vdash \Delta, \chi \\ | \\ D_2' \quad | \\ \Gamma' \vdash \Delta', \psi \end{array} \quad \begin{array}{c} D_2'' \\ \chi, \Gamma'' \vdash \Delta'' \end{array}}{\psi, \Gamma, \Gamma' \vdash \Delta, \Delta', \chi} \quad (\text{Cut}) \quad |$$

$$\frac{}{\psi, \Gamma, \Gamma' \vdash \Delta, \Delta', \chi} \quad (\text{Cut})$$

$$\Gamma, \Gamma', \Gamma'' \vdash \Delta, \Delta', \Delta''$$

**Complexity:** Two new cuts of rank  $< r$

**Case E:  $\phi = \forall x \psi(x)$**

$D_1$  ends with  $(\forall\text{-R})$ :

$$\frac{\begin{array}{c} D_1': \\ \Gamma \vdash \Delta, \psi(y) \text{ [y fresh]} \end{array}}{\Gamma \vdash \Delta, \forall x \psi(x)} (\forall\text{-R})$$

$D_2$  ends with  $(\forall\text{-L})$ :

$$\frac{\begin{array}{c} D_2': \\ \psi(t), \Gamma' \vdash \Delta' \text{ [t term]} \end{array}}{\forall x \psi(x), \Gamma' \vdash \Delta'} (\forall\text{-L})$$

**Reduction:**

$$\frac{\begin{array}{c} D_1'[y := t] \\ \Gamma \vdash \Delta, \psi(t) \text{ [substitute y with t]} \\ D_2' \\ \psi(t), \Gamma' \vdash \Delta' \end{array}}{\Gamma, \Gamma' \vdash \Delta, \Delta'} (\text{Cut on } \psi(t))$$

**Subtlety:** Substitution  $y := t$  may require  $\alpha$ -renaming

**Complexity:** One new cut of rank  $< r$

**Case F:  $\phi = \exists x \psi(x)$**

Symmetric to universal case.

### Phase 3: Multiple Cuts

**Problem:** If  $\phi$  appears multiple times:

$D_1$  produces:  $\Gamma \vdash \Delta, \phi, \phi, \phi$

$D_2$  consumes:  $\phi, \Gamma' \vdash \Delta'$

**Solution:** Duplication of  $D_2$

$$\begin{array}{c}
\Gamma \vdash \Delta, \varphi, \varphi, \varphi \\
\begin{array}{ccc}
| & | & | \\
D_2 & D_2' & D_2'' \\
\varphi, \Gamma' \vdash \Delta' & \varphi, \Gamma' \vdash \Delta' & \varphi, \Gamma' \vdash \Delta'
\end{array} \\
\hline
\Gamma, \Gamma', \Gamma', \Gamma' \vdash \Delta, \Delta', \Delta', \Delta'
\end{array}$$

**Explosion:** If  $k$  occurrences, size grows by factor  $k$

## 5.4 Precise Complexity Bounds

**Theorem (Gentzen, 1936):** If  $D$  has height  $h$  and cut-rank  $r$ , there exists  $D'$  cut-free with:

$$h(D') \leq H_r(h)$$

where  $H_r$  is the Ackermann hierarchy:

- $H_0(n) = n + 1$
- $H_{r+1}(n) = H_r^{(n+1)}(n)$  [iterate  $H_r$  for  $n+1$  times]

### Concrete Examples

For  $h = 100$ ,  $r = 3$ :

$$H_1(100) \approx 2 \cdot 100 = 200$$

$$H_2(100) \approx 2^{100}$$

$$H_3(100) \approx 2^{(2^{(2^{100}))})} \approx 10^{(10^{30})} \text{ (beyond atoms in universe)}$$

**Theorem (Orevkov, 1993)** - Optimal bounds: There exist derivations  $D_n$  such that:

1.  $h(D_n) = O(n)$
2.  $cr(D_n) = 3$
3. Every cut-free  $D'_n$  has  $h(D'_n) \geq H_3(n)$

**Consequence:** Gentzen's bounds are optimal!

**Corollary:** Cut-elimination is **non-elementary**—cannot be bounded by any finite tower of exponentials.

## 5.5 Ordinal Analysis of Elimination

**Ordinal assignment:**

For derivation  $D$  with cut-rank  $r$ , assign  $o_r(D)$ :

**Base:**

$$o_r(\text{axiom}) = 0$$

**Logical rules (non-cut):**

$$o_r(D'/D) = \omega^*(o_r(D'))$$

$$o_r(D_1, D_2/D) = \omega^*(\max(o_r(D_1), o_r(D_2)))$$

**Cut rule** on  $\varphi$  with  $|\varphi| = s$ :

$$o_r(D_1, D_2/D) =$$

$$\omega^*(\max(o_r(D_1), o_r(D_2)) + 1) \quad \text{if } s = r$$

$$\omega^*(\max(o_r(D_1), o_r(D_2))) \quad \text{if } s < r$$

**Theorem:** During cut-elimination:

- If  $cr(D) = r$  and we apply a reduction, we get  $D'$  with  $o_r(D') < o_r(D)$
- Therefore elimination terminates by  $TI(\epsilon_0)$ !

**Connection to Gentzen:** This explains why  $|PA| = \epsilon_0$ : cut-elimination for PA requires transfinite induction up to  $\epsilon_0$ .

## 5.6 Optimization Strategies

### Strategy 1: Bottom-Up Elimination

Eliminate cuts lowest in the derivation first.

**Advantage:** Reduces duplications **Disadvantage:** Doesn't always minimize height

### Strategy 2: Elimination by Degree

Eliminate cuts of smaller degree first.

**Advantage:** Optimizes final degree **Disadvantage:** Height can explode earlier

### Strategy 3: Stratification

Divide derivation into layers and proceed by levels.

**Advantage:** More controllable **Disadvantage:** Complex to implement

### Strategy 4: Normalization by Evaluation

Interpret derivations as programs and "execute".

**Advantage:** Connection with lambda-calculus **Disadvantage:** Requires sophisticated semantics

## 5.7 Cut-Elimination for Non-Classical Logics

### Intuitionistic Logic (LJ)

**Restriction:**  $\Delta$  contains at most 1 formula

**Modifications:**

- Disjunction case simpler (no duplication needed)
- Implication case unchanged
- Termination: still guaranteed with same ordinal  $\varepsilon_0$

### Linear Logic (LL)

**Complexity:** Much greater!

**Structure:**

- Formulas have "modalities" (!, ?)
- Contraction limited
- Procedure requires resource tracking

**Ordinal:** Depends on fragment, can reach  $\Gamma_0$  for complex fragments

### Modal Logic (K, S4, S5)

**Extension:** Rules for  $\Box$ ,  $\Diamond$

**Cut-elimination:**

- **K:** Similar to classical, ordinal  $\omega^\omega$
- **S4:** Complicated by transitivity, ordinal  $\varepsilon_0$
- **S5:** Simpler due to symmetry, ordinal  $\omega^\omega$

## 5.8 Multicut and Mix

**Multicut:**

$$\frac{\Gamma \vdash \Delta, \varphi_1, \dots, \varphi_n \quad \varphi_1, \dots, \varphi_n, \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'} \text{ (Multicut)}$$

**Theorem:** Multicut is eliminable, but with worse bounds:

$h(D') \leq H_r^n(h)$  [iterate  $H_r$   $n$  times]

**Mix rule:**

$$\frac{\Gamma \vdash \Delta \quad \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'} \text{ (Mix)}$$

**Surprise:** Mix is NOT always eliminable!

- Requires special properties of the sequent
- Eliminability equivalent to deep combinatorial properties

## 5.9 Cut-Elimination and Interpolation

**Craig's Theorem:** If  $\vdash \phi \rightarrow \psi$ , then there exists  $\chi$  (interpolant) such that:

1.  $\text{Voc}(\chi) \subseteq \text{Voc}(\phi) \cap \text{Voc}(\psi)$
2.  $\vdash \phi \rightarrow \chi$
3.  $\vdash \chi \rightarrow \psi$

**Proof via cut-elimination:**

1. Given derivation of  $\phi \vdash \psi$
2. Apply cut-elimination
3. Construct  $\chi$  from intermediate sequents
4.  $\chi$  emerges "naturally" from cut-free structure

**Application:** Fundamental in model theory and computability theory.

## 5.10 Fundamental Limits

**Theorem (Statman, 1979):** The problem "Does  $D$  have a cut-free derivation?" is **EXSPACE-complete**.

**Theorem (Beckmann-Buss, 2001):** There exist families  $\{\phi_n\}$  such that:

- $\phi_n$  has proof with cuts of length  $O(n)$
- Every cut-free proof of  $\phi_n$  has length  $\geq 2^{2^{(n/\log n)}}$

**Interpretation:** Cuts are **essential** for efficient proofs, even if theoretically eliminable.

**Philosophical consequence:** Even if cuts are "eliminable in theory", they are "indispensable in practice" for human mathematics.

## 5.11 Modern Applications

## 1. Proof Complexity

- Lower bounds for Resolution
- P vs NP approaches (proof-theoretic)

## 2. Program Extraction

- Curry-Howard: eliminate cuts = optimize code
- Compilers based on cut-elimination

## 3. Automated Theorem Proving

- **Hybrid strategy:** Search with cuts (fast), then normalize (certified)
- Resolution, Tableaux, Sequent calculus provers

## 4. Formal Verification

- Coq, Isabelle: produce cut-free proof terms
- Correctness guaranteed by normalization

## 5.12 Open Problems

1. **Optimal bounds:** Strategies that improve  $H_r(n)$ ?
2. **Specific logics:** Cut-elimination for temporal, description logics?
3. **Automated optimization:** Minimize height and degree simultaneously?
4. **Semantic methods:** More conceptual characterizations?
5. **Quantum logics:** Extension to quantum reasoning?

# Part 6: Interdisciplinary Applications and Connections

## 6.1 Computability Theory: The Hyperarithmetical Hierarchy

**Definition:** The hyperarithmetical hierarchy extends the arithmetical hierarchy beyond  $\omega$  using transfinite recursion.

For  $\alpha$  a recursive ordinal:

$H_\alpha$  ( $\alpha$ -recursive sets):

- $H_0$  = recursive sets
- $H_{\alpha+1} = \{A : A \leq_T H_{\alpha^{(\infty)}}\}$  (Turing jump)
- $H_\lambda = \bigcup_{\alpha < \lambda} H_\alpha$  for limit  $\lambda$

**Theorem (Kleene, 1955):**  $\bigcup_{\alpha < \omega_1^{\text{CK}}} H_\alpha = \Delta^1_1$  sets

where  $\omega_1^{\text{CK}}$  is the Church-Kleene ordinal.

## Proof-Theoretic Connection

| System                | Ordinal                          | Hierarchy                                          |
|-----------------------|----------------------------------|----------------------------------------------------|
| PA                    | $\varepsilon_0$                  | $H_{\varepsilon_0}$ contains all PA-definable sets |
| $\text{ACA}_0$        | $\varepsilon_0$                  | Comprehension limited to arithmetical hierarchy    |
| $\text{ATR}_0$        | $\Gamma_0$                       | Access to $H_\alpha$ for $\alpha < \Gamma_0$       |
| $\Pi^1_1\text{-CA}_0$ | $\psi_0(\varepsilon_{\Omega+1})$ | Entire hyperarithmetical hierarchy                 |

## 6.2 Program Termination

**Problem:** Given a program  $P$ , does it terminate on all inputs?

**Ordinal approach:** Assign to  $P$  an ordinal  $\alpha(P)$  that measures "termination complexity".

### Ackermann Function and Fast-Growing Hierarchy

For ordinal  $\alpha$ , define  $f_\alpha: \mathbb{N} \rightarrow \mathbb{N}$ :

- $f_0(n) = n + 1$
- $f_{\alpha+1}(n) = f_\alpha^n(n)$  (iterate  $n$  times)
- $f_\lambda(n) = f_{\lambda[n]}(n)$  where  $\lambda[n]$  is the  $n$ -th approximation of  $\lambda$

**Examples:**

- $f_\omega(n) \approx 2n$  (exponential)
- $f_{\omega+1}(n) \approx 2^{2^{\dots^2}}$  (tower of  $n$  exponentials)
- $f_{\omega \cdot 2}(n) = \text{"super-exponential"}$
- $f_{\varepsilon_0}(n) = \text{Ackermann function}$

**Theorem:** A program terminating in time  $< f_\alpha(n)$  can be verified in a system with proof-theoretic ordinal  $\alpha$ .

## Practical Applications

**Termination ordinals for common algorithms:**

- Quicksort:  $\omega^2$
- Heapsort:  $\omega \cdot \log n$



- Ackermann:  $\varepsilon_0$
- Goodstein sequences:  $\varepsilon_0$
- Hydra game:  $\varepsilon_0$

## 6.3 Type Theory and Natural Deduction Systems

**Extended Curry-Howard correspondence:**

| Logic                      | Type                                   | Ordinal                                  |
|----------------------------|----------------------------------------|------------------------------------------|
| $\perp$                    | Empty type                             | 0                                        |
| $\varphi \rightarrow \psi$ | $\varphi \rightarrow \psi$ (functions) | $\omega^\alpha \rightarrow \omega^\beta$ |
| $\forall x \varphi(x)$     | $\Pi x:A. B(x)$ (dependent)            | Limit                                    |
| Induction                  | W-types                                | $\varepsilon_0$                          |
| Transfinite recursion      | Induction-recursion                    | $> \varepsilon_0$                        |

### System F (Polymorphic Lambda Calculus)

**Properties:**

- Second-order polymorphism
- $|\text{System F}| = \varepsilon_0$
- Corresponds exactly to PA

### Calculus of Constructions

**Properties:**

- Full dependent types
- $|\text{CoC}| \gg \varepsilon_0$
- Foundation of Coq proof assistant

**Implication:** The logical strength of a type system is measured by its proof-theoretic ordinal.

## 6.4 Proof Mining: Extracting Bounds

**Kohlenbach's program:** Extract quantitative information from qualitative proofs.

**Example:** Banach Fixed Point Theorem

- **Classical proof:** "There exists  $x$  such that  $f(x) = x$ "
- **Proof mining:** "Given  $\varepsilon > 0$ , after  $n = \lceil \log(\varepsilon/d_0)/\log(c) \rceil$  iterations,  $|x_n - x| < \varepsilon$ "

### Technique

1. Translate to monotone higher-order logic
2. Apply functional interpretation (Dialectica)
3. Extract witnessing functionals
4. Compute concrete bounds

## Successes

### Nonlinear functional analysis:

- Fixed point theorems  $\rightarrow$  convergence rates
- Approximate solutions  $\rightarrow$  concrete algorithms

### Approximation theory:

- Existence theorems  $\rightarrow$  construction methods
- Best approximation  $\rightarrow$  computable bounds

### Ergodic theory:

- Ergodic theorems  $\rightarrow$  rates of convergence
- Mean ergodic theorem  $\rightarrow$  effective bounds

## 6.5 Large Cardinals and Large Ordinals

**Surprising connection:** Some large cardinals have "ordinal reflections" in proof theory.

| Large Cardinal           | Proof-Theoretic Ordinal          | System                           |
|--------------------------|----------------------------------|----------------------------------|
| Inaccessible             | Limit of predicative systems     | Beyond $\Gamma_0$                |
| Mahlo                    | $\psi_\Omega(\varepsilon_{M+1})$ | $ID_{<\omega}$                   |
| Weakly compact           | $\psi(K)$                        | KPM                              |
| $\Pi^1_1$ -indescribable | —                                | $\Pi^1_1$ -CA <sub>0</sub> + ... |

**Theorem (Rathjen):** The proof strength of "ZFC +  $\exists$  Mahlo cardinal" corresponds to:  $\psi_\Omega(\varepsilon_{M+1})$

where M is a symbol for "first Mahlo".

## Philosophical Significance

Large cardinals in set theory have "echoes" in ordinal notation systems, suggesting deep structural connections between:

- Set-theoretic strength
- Proof-theoretic ordinals
- Computational complexity

## 6.6 Category Theory: Arithmetic Topoi

**Effective topos Eff:** Category of "sets with computable structure"

- Objects: Realizable sets
- Morphisms: Realizable functions

**Properties:**

- Eff models intuitionistic logic
- Natural numbers object has proof-theoretic strength  $\varepsilon_0$
- Subtopoi correspond to induction restrictions

**Sheaf topoi on ordinals:**

- $\text{Sh}(\alpha)$  = sheaves on topology of ordinals  $< \alpha$
- $|\text{Sh}(\alpha)| \leq \alpha$  (in proof-theoretic sense)

### Applications

**Realizability theory:** Different notions of realizability correspond to different ordinal levels:

- Kleene realizability: Basic recursion ( $\omega^\omega$ )
- Modified realizability: Extended systems
- Dialectica categories: Full hierarchies

## 6.7 Nonstandard Analysis: Ultrapowers

**Robinson's construction:**

- ${}^*\mathbb{N} = \mathbb{N}^\mathbb{N}/U$  where  $U$  is a non-principal ultrafilter
- Contains "nonstandard infinities"

### Ordinal Connection

For every  $\alpha < \varepsilon_0$ , there exists  $x \in {}^*\mathbb{N}$  such that:

- $x$  is "as large as"  $f_\alpha$  for some fundamental sequence

**Transfer principle:** Truth in  $\mathbb{N} \approx$  Truth in  ${}^*\mathbb{N}$  for bounded formulas

**Proof strength:** NSA (nonstandard analysis)  $\approx$   $\text{ACA}_0 + \text{WKL}_0$

## 6.8 Infinite Games and Determinacy

**Gale-Stewart games:**

- Two players alternately choose  $n_0, n_1, n_2, \dots$

- Winning set  $A \subseteq \mathbb{N}^\omega$
- Player I wins if  $(n_0, n_1, \dots) \in A$

**Axiom of Determinacy (AD):** Every Gale-Stewart game is determined.

## Borel Hierarchy

- $\Sigma^0_n, \Pi^0_n$  = Borel sets of class  $n$
- **Theorem (Martin):**  $ZFC \vdash \text{Det}(\Sigma^0_n)$  for all  $n$

**Proof-theoretic connection:**

| Determinacy                    | Strength                                        |
|--------------------------------|-------------------------------------------------|
| $\text{Det}(\text{open sets})$ | $\text{ATR}_0$                                  |
| $\text{Det}(\Sigma^0_3)$       | $\Pi^1_1\text{-CA}_0$                           |
| $\text{Det}(\text{Borel})$     | $ZFC + \exists \omega \text{ Woodin cardinals}$ |

**Ordinals:** The ordinal of Borel determinacy is beyond any standard constructive ordinal notation.

## 6.9 Descriptive Set Theory

**Projective hierarchy:**

- $\Sigma^1_n$  = projection of  $\Pi^1_{n-1}$
- $\Pi^1_n$  = complement of  $\Sigma^1_n$
- $\Delta^1_n = \Sigma^1_n \cap \Pi^1_n$

### Proof Strength

| Principle                   | System                               | Ordinal                                 |
|-----------------------------|--------------------------------------|-----------------------------------------|
| $\Delta^1_1$ -comprehension | $\text{ATR}_0$                       | $\Gamma_0$                              |
| $\Pi^1_1$ -comprehension    | $\Pi^1_1\text{-CA}_0$                | $\psi_0(\varepsilon_{\Omega+1})$        |
| $\Delta^1_2$ -comprehension | $\Delta^1_2\text{-CA}_0 + \text{BI}$ | $\psi_0(\varepsilon_{\Omega_\omega+1})$ |

## 6.10 Philosophy of Mathematics: Predicativism

**Predicative program** (Weyl, Feferman):

- Avoid vicious circles in definitions
- Allow quantification only over "already constructed" objects

**Fundamental result (Feferman-Schütte):**

- Predicative mathematics reaches exactly up to  $\Gamma_0$
- Beyond  $\Gamma_0$  requires genuine impredicativity

**Interpretation**

- Ordinals  $< \Gamma_0$  = "generalized finite"
- Ordinals  $\geq \Gamma_0$  = "essential transfinite"

**Contemporary debate:**

- Is  $\Gamma_0$  the natural boundary of "constructive" mathematics?
- $\text{ATR}_0$  is the maximal predicative system + transfinite recursion
- Some schools accept  $\Pi^1_1\text{-CA}_0$  as still predicative

## 6.11 Contemporary Applications

### 1. Formal Software Verification

**Proof assistants** (Coq, Agda, Lean) based on ordinals:

- Termination checking uses ordinal assignments
- Certification of critical programs (avionics, medical devices)
- Proof-theoretic ordinal determines expressiveness

### 2. Computational Complexity

**Correspondence:**

- $\text{PTIME} \leftrightarrow f_{\omega}$
- $\text{EXPTIME} \leftrightarrow f_{\omega^2}$
- Elementary functions  $\leftrightarrow f_{\epsilon_0}$
- Primitive recursive  $\leftrightarrow f_{\Gamma_0}$

### 3. Theoretical Machine Learning

**Neural network depth**  $\leftrightarrow$  Ordinals:

- Shallow networks: finite ordinals
- Deep networks:  $\omega, \omega^2$
- Recursive architectures: beyond  $\omega^\omega$

**Expressivity:** Higher ordinal = greater expressivity

### 4. String Theory / Theoretical Physics

**Conjectures:**

- Classification of backgrounds using ordinal hierarchies
- Computational limits of the universe
- Connection to information-theoretic bounds

## 6.12 Open Problems

1. **Ordinal of ZFC:** Still unknown precisely (only bounds available)
2. **Large cardinals and ordinals:** Complete connection still mysterious
3. **Proof complexity:** Relation between ordinals and lower bounds for proofs
4. **Non-well-founded systems:** Extend ordinal theory to sets with circularity
5. **Generalized predicative ordinals:** Beyond  $\Gamma_0$ , what is the next "plateau"?
6. **Quantum proof theory:** Do quantum logics have ordinal measures?
7. **Homotopy type theory:** What replaces ordinals in univalent foundations?

## 6.13 Unified Picture

All these connections suggest a deep unifying structure:

**Ordinals measure:**

- Logical strength (proof theory)
- Computational complexity (recursion theory)
- Set-theoretic size (large cardinals)
- Constructive content (predicativity)
- Type-theoretic power (dependent types)

This is not coincidence but reflects fundamental properties of **hierarchical structure** in mathematics.

# Part 7: Cut-Elimination - Deep Technical Analysis

## 7.1 The Cut Problem: Deep Motivation

**Philosophical intuition:** The cut represents the use of intermediate lemmas in a proof.

**Scenario:** I want to prove  $\Gamma \vdash \Delta$

**Strategy:** Find a lemma  $\varphi$  such that:

1.  $\Gamma \vdash \varphi$  (prove the lemma)
2.  $\varphi \vdash \Delta$  (use the lemma)

### Epistemological Problem

- $\varphi$  can be arbitrarily complex
- $\varphi$  may have no "obvious" relation to  $\Gamma$  or  $\Delta$
- How do we know the lemma is "necessary"?

**Gentzen's answer:** The lemma is NOT necessary—every proof can be made **analytic** (using only subformulas of the conclusion).

## 7.2 Complexity Measures for Derivations

For a derivation  $D$ , we define several measures:

### Height $h(D)$

Length of the longest branch

### Degree $g(D)$

Maximum logical complexity of formulas

### Cut-rank $cr(D)$

Maximum complexity of cut formulas

### Multiplicity $m(D)$

Total number of cut applications

### Trade-off theorem:

- Eliminating cuts can explode  $h(D)$
- But  $g(D)$  and  $cr(D)$  reduce to 0
- No procedure minimizes all measures simultaneously

## 7.3 Elimination Procedure: Detailed Cases

Consider a cut on formula  $\varphi$ :

$$\begin{array}{c}
 \begin{array}{cc}
 D_1 & D_2 \\
 \hline
 \Gamma \vdash \Delta, \varphi & \varphi, \Gamma' \vdash \Delta'
 \end{array} \\
 \hline
 \Gamma, \Gamma' \vdash \Delta, \Delta' \quad \text{(Cut)}
 \end{array}$$

### Case 1: $\varphi$ is Atomic and Axiomatic

If  $\varphi \in \Gamma$  and  $\varphi \in \Delta$ , the cut reduces to:

$$\frac{\Gamma \vdash \Delta, \varphi \quad \varphi, \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'} \rightarrow \Gamma, \Gamma' \vdash \Delta, \Delta' \text{ (axiom)}$$

**Complexity:**  $O(1)$  - immediate elimination

### Case 2: $\varphi = \psi \wedge \chi$ Introduced by $(\wedge\text{-R})$ and $(\wedge\text{-L})$

**D<sub>1</sub> structure:**

$$\frac{\Gamma \vdash \Delta, \psi \quad \Gamma \vdash \Delta, \chi}{\Gamma \vdash \Delta, \psi \wedge \chi} (\wedge\text{-R})$$

**D<sub>2</sub> structure** (if using  $\wedge\text{-L}_1$ ):

$$\frac{\psi, \chi, \Gamma' \vdash \Delta'}{\psi \wedge \chi, \Gamma' \vdash \Delta'} (\wedge\text{-L}_1)$$

**Reduction:**

$$\begin{array}{c} \frac{\Gamma \vdash \Delta, \psi \quad \psi, \chi, \Gamma' \vdash \Delta'}{\Gamma, \chi, \Gamma' \vdash \Delta, \Delta'} \text{ (Cut on } \psi, \text{ rank} < \varphi) \\ \Gamma \vdash \Delta, \chi \\ \hline \Gamma, \Gamma', \Gamma' \vdash \Delta, \Delta', \Delta' \text{ (Cut on } \chi, \text{ rank} < \varphi) \\ \hline \Gamma, \Gamma' \vdash \Delta, \Delta' \text{ (Contractions)} \end{array}$$

**Crucial observation:** New cuts are on  $\psi$  and  $\chi$ , which are proper subformulas of  $\varphi$ .

### Case 3: $\varphi = \forall x \psi(x)$ with Quantifier Rules

**D<sub>1</sub>:**

$$\frac{\Gamma \vdash \Delta, \psi(y) \text{ [y fresh]}}{\Gamma \vdash \Delta, \forall x \psi(x)} (\forall\text{-R})$$



$$\Gamma \vdash \Delta, \forall x \psi(x)$$

**D<sub>2</sub>:**

$$\frac{\psi(t), \Gamma' \vdash \Delta' \text{ [t term]}}{\forall x \psi(x), \Gamma' \vdash \Delta'} \quad (\forall\text{-L})$$

**Reduction:**

$$\frac{\frac{\Gamma \vdash \Delta, \psi(y)}{\Gamma \vdash \Delta, \psi(t)} \text{ [substitute y with t]}}{\frac{\psi(t), \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'}} \quad (\text{Cut on } \psi(t), \text{rank} < \varphi)$$

**Subtlety:** Substitution  $y := t$  may require  $\alpha$ -renaming

### Case 4: Multiple Cuts (The Hard Case)

If  $\varphi$  appears multiple times:

$D_1$  produces:  $\Gamma \vdash \Delta, \varphi, \varphi, \varphi$

$D_2$  consumes:  $\varphi, \Gamma' \vdash \Delta'$

Must duplicate  $D_2$  for each occurrence of  $\varphi$ :

$$\frac{\begin{array}{c} \Gamma \vdash \Delta, \varphi, \varphi, \varphi \\ \varphi, \Gamma' \vdash \Delta' \quad \varphi, \Gamma' \vdash \Delta' \quad \varphi, \Gamma' \vdash \Delta' \end{array}}{\Gamma, \Gamma', \Gamma', \Gamma' \vdash \Delta, \Delta', \Delta', \Delta'}$$

**Exponential explosion:** If  $m(D)$  = number of cuts, elimination can produce derivation of size  $2^{(2^{(2^{(2^{\dots})})})}$  (tower of exponentials).

## 7.4 Precise Complexity Bounds

**Theorem (Orevkov, 1993):**

Let  $D$  be a derivation of length  $n$  with cut-rank  $r$ . There exists a cut-free derivation  $D'$  of length:

$$h(D') \leq H_r(n)$$

where  $H_r$  is the extended Ackermann hierarchy:

- $H_0(n) = n + 1$
- $H_{r+1}(n) = H_r^{(n+1)}(n)$  [iterate  $H_r$  for  $n+1$  times]

## Concrete Examples

For  $n = 100$ ,  $r = 3$ :

- $H_1(100) \approx 2 \cdot 100 = 200$
- $H_2(100) \approx 2^{100}$
- $H_3(100) \approx 2^{(2^{(2^{100}))})} \approx 10^{(10^{30})}$  (more than atoms in universe)

**Philosophical consequence:** Cut-free proofs are:

- Conceptually simple (analytic)
- Practically impossible to write
- Testament to the necessity of lemmas!

**Theorem (Orevkov):** These bounds are **optimal**!

**Corollary:** Cut-elimination is **non-elementary**—cannot be bounded by any finite tower of exponentials.

## 7.5 Ordinal Analysis of Elimination

**Refined ordinal assignment:**

For derivation  $D$  with cut-rank  $r$ , define  $o_r(D)$ :

**Base:**

$$o_r(\text{axiom}) = 0$$

**Logical rules** (non-cut):

$$o_r(D'/D) = \omega^{(o_r(D'))}$$

$$o_r(D_1, D_2/D) = \omega^{(\max(o_r(D_1), o_r(D_2)))}$$

**Cut rule** on  $\varphi$  with  $|\varphi| = s$ :

$$o_r(D_1, D_2/D) =$$

$$\omega^{(\max(o_r(D_1), o_r(D_2)) + 1)} \quad \text{if } s = r$$

$$\omega^{(\max(o_r(D_1), o_r(D_2)))} \quad \text{if } s < r$$

**Theorem:** During cut-elimination:

- If  $cr(D) = r$  and we apply a reduction, we get  $D'$  with  $o_r(D') < o_r(D)$
- Therefore elimination terminates by  $TI(\epsilon_0)$ !

**Connection to Gentzen:** This explains why  $|PA| = \epsilon_0$ : cut-elimination for PA requires transfinite induction up to  $\epsilon_0$ .

## Key Property

$o_r(D) < \epsilon_0$  if  $r \leq 1$  (propositional cuts)

For larger  $r$ ,  $o_r(D)$  can grow beyond  $\epsilon_0$ .

**Theorem (Buchholz):** For PA, every derivation of length  $n$  has:  $o(D) < \epsilon_0 \cdot \omega^\omega \dots \omega$  ( $n$  times)

## 7.6 Structural Variations of Cut-Elimination

### Multicut (Simultaneous Cut on Multiple Formulas)

$$\frac{\Gamma \vdash \Delta, \varphi_1, \dots, \varphi_n \quad \varphi_1, \dots, \varphi_n, \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'} \text{ (Multicut)}$$

**Theorem:** Multicut is eliminable, but with even worse bounds:  $h(D') \leq H_r^n(h)$  [iterate  $H_r$   $n$  times]

### Mix Rule (Merging Sequents)

$$\frac{\Gamma \vdash \Delta \quad \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'} \text{ (Mix)}$$

**Surprise:** Mix is NOT always eliminable!

- Requires special properties of the sequent
- Eliminability equivalent to deep combinatorial properties

## 7.7 Gentzen's Hauptsatz for Non-Classical Logics

### Intuitionistic Logic (LJ)

**Restriction:**  $\Delta$  contains at most 1 formula

**Modifications:**

- Disjunction case simpler (no duplication needed)
- Implication case unchanged

- Termination: still guaranteed with same ordinal  $\varepsilon_0$

## Linear Logic (LL)

**Complexity:** Much greater!

**Structure:**

- Formulas have "modalities" ( $!$ ,  $?$ )
- Contraction limited
- Procedure requires resource tracking

**Ordinal:** Depends on fragment, can reach  $\Gamma_0$  for complex fragments

## Modal Logic (K, S4, S5)

**Extension:** Rules for  $\Box$ ,  $\Diamond$

**Cut-elimination:**

- **K:** Similar to classical, ordinal  $\omega^\omega$
- **S4:** Complicated by transitivity, ordinal  $\varepsilon_0$
- **S5:** Simpler due to symmetry, ordinal  $\omega^\omega$  (easier!)

## 7.8 Cut-Elimination as Normalization

**Curry-Howard correspondence:**

| Sequents    | $\lambda$ -Calculus             | Description     |
|-------------|---------------------------------|-----------------|
| Derivation  | $\lambda$ -term                 | Program         |
| Cut         | $\beta$ -redex $(\lambda x.M)N$ | Application     |
| Cut-free    | Normal form                     | Reduced program |
| Elimination | $\beta$ -reduction              | Execution       |

**Church-Rosser Theorem:**

- Order of cut reduction doesn't matter
- All maximal reductions lead to same normal derivation

**Computational consequence:**

- Eliminating cuts = executing a program
- Cut-free derivations = values
- Non-termination =  $\perp$  derivable (inconsistency)

## 7.9 Fundamental Limits

**Theorem (Statman, 1979):** The problem "Does  $D$  have a cut-free derivation?" is **EXPSpace-complete**.

**Theorem (Beckmann-Buss, 2001):** There exist families of theorems  $\{\varphi_n\}$  such that:

- $\varphi_n$  has proof with cuts of length  $O(n)$
- Every cut-free proof of  $\varphi_n$  has length  $\geq 2^{2^{n/\log n}}$

**Interpretation:** Cuts are **essential** for efficient proofs, even if theoretically eliminable.

**Philosophical consequence:** Even though cuts are "eliminable in theory", they are "indispensable in practice" for human mathematics.

## 7.10 Modern Applications

### 1. Proof Complexity

**Lower bounds for Resolution:**

- Resolution = cut-elimination for clauses
- Exponential lower bounds via cut complexity

**P vs NP approach** (proof-theoretic):

- If  $P \neq NP$ , certain theorems require large cuts
- Cut-elimination complexity relates to circuit complexity

### 2. Program Extraction

**Curry-Howard:**

- Eliminate cuts = optimize code
- Compilers based on cut-elimination
- Remove intermediate computations (inlining)

**Paradox:** Eliminating cuts makes programs longer but not necessarily faster!

### 3. Automated Theorem Proving

**Hybrid strategy:**

1. Search for proof with cuts (much faster)
2. If found, normalize to produce certificate (slower but certified)

**Implementations:**

- **Resolution:** Essentially cut-elimination for clauses

- **Tableaux:** Avoids cuts by building models
- **Sequent calculus provers:** Implement Gentzen directly

**Trade-off:** Speed vs. certification

## 4. Formal Verification

**Proof assistants** (Coq, Isabelle):

- Produce cut-free proof terms
- Correctness guaranteed by normalization
- But internal proofs use cuts for efficiency

## 7.11 Contemporary Generalizations

### Deep Inference

Allows rule application "deep" within formulas.

**Advantage:** More expressive **Disadvantage:** Cut-elimination much more complex

### Nested Sequents

Nested sequents for modal logics.

**Structure:**  $\Gamma \vdash \Delta [\Gamma_1 \vdash \Delta_1 [\Gamma_2 \vdash \Delta_2]]$

**Cut-elimination:** Possible but requires new techniques

### Focused Proof Systems

Restricts rules to optimize search.

**Property:** Cut-elimination more efficient in practice

### Infinitary Logic

**Derivations** can be infinite ( $\omega$ -rule)

**Ordinals:** Transfinite ordinals necessary

**Applications:** Descriptive set theory

## 7.12 Cut-Elimination and Interpolation

**Craig's Theorem:** If  $\vdash \phi \rightarrow \psi$ , then there exists  $\chi$  (interpolant) such that:

1.  $\text{Voc}(\chi) \subseteq \text{Voc}(\phi) \cap \text{Voc}(\psi)$

2.  $\vdash \varphi \rightarrow \chi$
3.  $\vdash \chi \rightarrow \psi$

**Proof via cut-elimination:**

1. Given derivation of  $\varphi \vdash \psi$
2. Apply cut-elimination
3. Construct  $\chi$  from intermediate sequents
4.  $\chi$  emerges "naturally" from cut-free structure

**Application:** Fundamental in model theory and computability theory.

## 7.13 Semantic Cut-Elimination

**Alternative idea** (Schütte): "Semantic" elimination

**Method:**

1. Interpret derivations in ordinal structures
2. Cut is "non-constructive jump"
3. Elimination = replacement with explicit construction

**Advantage:** More conceptual, direct connection to ordinals **Disadvantage:** Less algorithmic

## 7.14 Open Problems

1. **Optimal bounds:** Are there strategies that improve  $H_r(n)$ ?
2. **Complexity of specific logics:** Cut-elimination for temporal, description, epistemic logics?
3. **Automated optimization:** Algorithms that minimize height and degree simultaneously?
4. **Semantic methods:** More conceptual characterizations via model theory?
5. **Quantum cut-elimination:** Extension to quantum logics?
6. **Machine learning:** Can neural networks learn optimal cut-elimination strategies?

## Part 8: Church-Kleene Ordinal $\omega_1^{CK}$

### 8.1 Recursive Ordinals: Rigorous Definition

**Fundamental definition:** An ordinal  $\alpha$  is **recursive** if there exists a recursive well-ordering of  $\mathbb{N}$  of order type  $\alpha$ .

**Formally:**  $\alpha$  is recursive  $\Leftrightarrow \exists R \subseteq \mathbb{N}^2$  such that:

1.  $R$  is recursive (decidable)
2.  $(\mathbb{N}, R)$  is a well-ordering
3.  $\text{order\_type}(\mathbb{N}, R) = \alpha$

## Key Properties

**Examples of recursive ordinals:**

- Every finite ordinal is recursive
- $\omega$  is recursive (natural ordering of  $\mathbb{N}$ )
- $\omega + 1, \omega \cdot 2, \omega^2$  are recursive
- $\varepsilon_0$  is recursive (via Cantor normal form representation)

**Theorem (Closure):** The class of recursive ordinals is closed under:

- Ordinal addition:  $\alpha, \beta$  recursive  $\Rightarrow \alpha + \beta$  recursive
- Ordinal multiplication:  $\alpha, \beta$  recursive  $\Rightarrow \alpha \cdot \beta$  recursive
- Exponentiation:  $\alpha$  recursive  $\Rightarrow \omega^\alpha$  recursive
- Countable limit: if  $\{\alpha_i\}$  is a recursive sequence of recursive ordinals, then  $\sup\{\alpha_i\}$  may be recursive (under conditions)

## 8.2 The Church-Kleene Ordinal $\omega_1^{\text{CK}}$

**Definition:**  $\omega_1^{\text{CK}}$  is the supremum of all recursive ordinals:

$$\omega_1^{\text{CK}} = \sup\{\alpha : \alpha \text{ is a recursive ordinal}\}$$

### Fundamental Theorem (Church-Kleene, 1938)

$\omega_1^{\text{CK}}$  is well-defined and:

1.  $\omega_1^{\text{CK}}$  itself is NOT recursive
2.  $\omega_1^{\text{CK}}$  is countable (as a set)
3.  $\omega_1^{\text{CK}}$  is the first non-recursive ordinal

## Equivalent Characterizations

**Via computability theory:**  $\omega_1^{\text{CK}} = \min\{\alpha : \alpha \text{ has no recursive notation}\}$

**Via arithmetical hierarchy:**  $\omega_1^{\text{CK}}$  = length of the hyperarithmetical hierarchy

**Via model theory:**  $\omega_1^{\text{CK}}$  = height of the smallest admissible model of Kripke-Platek



**Via infinitary logic:**  $\omega_1^{\text{CK}}$  = first height where  $L_{\omega_1 \omega}$  becomes "too large" for recursive semantics

## 8.3 The Hyperarithmetical Hierarchy

The hyperarithmetical hierarchy extends the arithmetical hierarchy beyond  $\omega$  using transfinite recursion.

### Definition by Levels

**Level 0:**  $H_0 = \{A \subseteq \mathbb{N} : A \text{ is recursive}\}$

**Successor level:**  $H_{\alpha+1}$  = set of all sets Turing-reducible to the jump of  $H_\alpha$ :

$$H_{\alpha+1} = \{A : A \leq_T H_\alpha^{(\infty)}\}$$

where  $H_\alpha^{(\infty)} = \{\langle e, n \rangle : \varphi_e^{H_\alpha}(n) \downarrow\}$  is the jump

**Limit level:**  $H_\lambda = \bigcup_{\alpha < \lambda} H_\alpha$  for limit ordinal  $\lambda$

**Complete extension:**  $\text{HYP} = \bigcup_{\alpha < \omega_1^{\text{CK}}} H_\alpha$  = complete hyperarithmetical hierarchy

### Kleene's Theorem (1955)

$$\text{HYP} = \Delta_1^1 = \{A \subseteq \mathbb{N} : A \text{ is both } \Sigma_1^1 \text{ and } \Pi_1^1\}$$

**Interpretation:** Hyperarithmetical sets are exactly the second-order arithmetical sets with minimal complexity.

## 8.4 Recursive Ordinal Notations

An **ordinal notation system** provides effective representations of ordinals.

### Kleene's System O

$O \subseteq \mathbb{N}$  is defined inductively:

1.  $0 \in O$  (represents ordinal 0)
2. If  $n \in O$ , then  $2^n \in O$  (represents  $|n|_O + 1$ )
3. If  $\{\varphi_e(i)\}_{i \in \mathbb{N}} \subseteq O$  is an increasing sequence, then  $3 \cdot 5^e \in O$  (represents  $\sup_i |\varphi_e(i)|_O$ )

### Ordinal Norm

$$|\cdot|_O : O \rightarrow \omega_1^{\text{CK}}:$$

- $|0|_O = 0$

- $|2^n|_O = |n|_O + 1$
- $|3 \cdot 5^e|_O = \sup_i |\varphi_e(i)|_O$

## Properties of O

- O is  $\Pi^1_1$ -complete (not hyperarithmetical!)
- $|O|_O = \{\alpha : \alpha < \omega_1^{CK}\} = \text{all recursive ordinals}$
- O provides notation for every ordinal  $< \omega_1^{CK}$

**Theorem (Spector, 1955):** O is the simplest  $\Pi^1_1$  set that is not hyperarithmetical.

## 8.5 Explicit Construction of $\omega_1^{CK}$

We can approximate  $\omega_1^{CK}$  through a constructive hierarchy:

### Fundamental Sequence

Define  $\alpha_n$  iteratively:

- $\alpha_0 = 0$
- $\alpha_{n+1} = \text{"first recursive ordinal not constructible with notations using only } \alpha_0, \dots, \alpha_n \text{"}$

### Explicit Levels

$\alpha_1 = \omega$  (natural order)

$\alpha_2 = \omega^2$  (lexicographic grid)

$\alpha_3 = \omega^\omega$  (finite sequences)

$\alpha_4 = \varepsilon_0$  (Cantor normal form)

$\alpha_5 = \Gamma_0$  (complete Veblen hierarchy)

...

**Theorem:**  $\omega_1^{CK} = \sup_{n \in \mathbb{N}} \alpha_n$  where each  $\alpha_n$  is effectively constructible but the sequence itself is not recursive.

## 8.6 $\omega_1^{CK}$ and Proof Theory

### Relation to Proof-Theoretic Ordinals

**Theorem:** For every recursively axiomatizable system S:

If  $|S|$  is the proof-theoretic ordinal of S, then  $|S| < \omega_1^{CK}$

### Examples

$|PA| = \varepsilon_0 < \omega_1^{CK}$

$|ATR_0| = \Gamma_0 < \omega_1^{CK}$

$$|\Pi^1_1\text{-CA}_0| = \psi_0(\varepsilon_{\Omega+1}) < \omega_1^{\text{CK}}$$

## Upper Bound Theorem

$$|\text{KP}| \approx \omega_1^{\text{CK}}$$

(approximation, technically slightly below)

Kripke-Platek set theory has proof-theoretic strength close to  $\omega_1^{\text{CK}}$ .

**Philosophical consequence:**  $\omega_1^{\text{CK}}$  marks the absolute limit of "explicit" proof-theoretic analysis—beyond this point, methods become fundamentally non-constructive.

## 8.7 Admissible Sets and $\omega_1^{\text{CK}}$

**Admissible set:** A structure  $(M, \in)$  satisfying the Kripke-Platek (KP) axioms.

### Barwise's Theorem

$$L_{\omega_1^{\text{CK}}} = \text{minimal admissible set containing } \mathbb{N}$$

where  $L_\alpha$  is Gödel's constructible hierarchy.

### Properties of $L_{\omega_1^{\text{CK}}}$

1.  $L_{\omega_1^{\text{CK}}} \models \text{KP}$
2.  $L_{\omega_1^{\text{CK}}}$  contains all hyperarithmetical sets
3.  $L_{\omega_1^{\text{CK}}} \cap \mathcal{P}(\mathbb{N}) = \text{HYP}$
4.  $\omega_1^{\text{CK}}$  is the first  $\alpha$  where  $L_\alpha$  is "closed" under recursive operations

## 8.8 Ordinal Indices and Complexity

**Definition:** For  $\alpha < \omega_1^{\text{CK}}$ , define:

$\text{deg}(\alpha)$  = minimal Turing complexity to encode  $\alpha$

$\text{deg}(\alpha) \in \{0', 0'', 0''', \dots\}$  (Turing jump degrees)

### Hierarchy Theorem

For every recursive ordinal  $\alpha$ :

- If  $\alpha < \omega^\omega$ , then  $\text{deg}(\alpha) = 0$  (recursive)
- If  $\omega^\omega \leq \alpha < \varepsilon_0$ , then  $\text{deg}(\alpha) \leq 0'$
- If  $\varepsilon_0 \leq \alpha < \Gamma_0$ , then  $\text{deg}(\alpha) \leq 0''$
- As  $\alpha \rightarrow \omega_1^{\text{CK}}$ ,  $\text{deg}(\alpha) \rightarrow \infty$  in the jump hierarchy

**Consequence:** Ordinals near  $\omega_1^{\text{CK}}$  require arbitrarily high computational complexity to be represented.

## 8.9 Relation to Second-Order Analysis

**Connection theorem:**

Levels of the hyperarithmetical hierarchy correspond to fragments of  $Z_2$ :

| Level                                       | System                | Ordinal                          |
|---------------------------------------------|-----------------------|----------------------------------|
| $H_0$                                       | $\text{RCA}_0$        | $\omega^\omega$                  |
| $H_{\{\varepsilon_0\}}$                     | $\text{ACA}_0$        | $\varepsilon_0$                  |
| $H_{\{\Gamma_0\}}$                          | $\text{ATR}_0$        | $\Gamma_0$                       |
| $\text{HYP} = H_{\{\omega_1^{\text{CK}}\}}$ | $\Pi^1_1\text{-CA}_0$ | $\psi_0(\varepsilon_{\Omega+1})$ |

**Key theorem:**  $A \in \text{HYP} \Leftrightarrow A$  is definable in  $Z_2$  by a  $\Delta^1_1$  formula

## 8.10 $\omega_1^{\text{CK}}$ in Infinitary Logic

**Logic  $L_{\{\omega_1, \omega\}}$ :** Allows countable conjunctions and disjunctions, but only finite quantifiers.

### Barwise's Theorem

A formula  $\varphi \in L_{\{\omega_1, \omega\}}$  is recursively decidable  $\Leftrightarrow \varphi$  has quantifier-rank  $< \omega_1^{\text{CK}}$

**Application:**  $\omega_1^{\text{CK}}$  measures the complexity of infinitary formulas that maintain decidability.

## 8.11 Beyond $\omega_1^{\text{CK}}$ : Projective Ordinals

**Generalization:** We can define higher ordinals using more complex hierarchies.

**Definition:**  $\omega_1^{\text{CK}}\{\Pi^1_1\} =$  supremum of ordinals with  $\Pi^1_1$  notation

### Hierarchy

$$\omega_1^{\text{CK}} = \omega_1^{\text{CK}}\{\Delta^1_1\}$$

$$\omega_1^{\text{CK}}\{\Sigma^1_1\} = \omega_1^{\text{CK}}\{\Pi^1_1\} = \omega_1^{\text{CK}} \text{ (same!)}$$

$$\omega_1^{\text{CK}}\{\Delta^1_2\} > \omega_1^{\text{CK}} \text{ (much larger)}$$

$$\omega_1^{\text{CK}}\{\Pi^1_2\} \text{ even larger}$$

...

**Theorem:** Under determinacy axioms, these ordinals have rich structure and are connected to proof-theoretic ordinals of very strong systems.

## 8.12 Computational Applications

### Program Termination

A program with ordinal complexity  $\alpha < \omega_1^{\text{CK}}$  has "effectively verifiable" termination.

**Examples:**

- Quicksort: ordinal  $\omega^2$
- Ackermann: ordinal  $\varepsilon_0$
- Hyperarithmetical algorithms: ordinal  $< \omega_1^{\text{CK}}$

**Theorem:** If an algorithm has complexity  $\geq \omega_1^{\text{CK}}$ , its termination is not verifiable by any recursive system.

## 8.13 Generalized Incompleteness Theorem

**Ordinal version of Gödel's theorem:**

For every recursively axiomatizable system  $S$  with  $|S| < \omega_1^{\text{CK}}$ :

1.  $S$  cannot prove  $\text{TI}(\omega_1^{\text{CK}})$
2.  $S$  cannot completely define its own model hierarchy
3. There exists a  $\Pi^1_1$  formula true but not provable in  $S$

**Interpretation:**  $\omega_1^{\text{CK}}$  is the "horizon of incompleteness" for recursive systems—it is the first ordinal completely inaccessible to recursive formalization.

## 8.14 Connection with Descriptive Set Theory

**Theorem (Suslin-Kleene):** For  $A \subseteq \mathbb{N}$ :

$$A \in \Delta^1_1 \Leftrightarrow A \in H_{\omega_1^{\text{CK}}}$$

$$A \text{ is } \Sigma^1_1 \Leftrightarrow A \text{ is a recursive enumeration in some } H_\alpha, \alpha < \omega_1^{\text{CK}}$$

**Effective Borel hierarchy:** The "effective" Borel hierarchy terminates exactly at level  $\omega_1^{\text{CK}}$ .

## 8.15 Open Problems about $\omega_1^{\text{CK}}$

1. **Exact value:** What is the "closed form" of  $\omega_1^{\text{CK}}$  in the large ordinal hierarchy?
2. **Distance from  $\Gamma_0$ :** How "far" is  $\omega_1^{\text{CK}}$  from  $\Gamma_0$ ? We know  $\Gamma_0 \ll \omega_1^{\text{CK}}$ , but precise bounds?

3. **Fine structure:** Is there a "fine hierarchy" between  $\Gamma_0$  and  $\omega_1^{\text{CK}}$  with natural structure?
4. **Generalizations:** How do the analogues  $\omega_\alpha^{\text{CK}}$  behave for  $\alpha > 1$ ?
5. **Determinacy:** Which set-theoretic assumptions determine exactly  $\omega_1^{\text{CK}}$ ?

## 8.16 Philosophical Significance

$\omega_1^{\text{CK}}$  represents several fundamental boundaries:

**As supremum of recursive ordinals:**

- Ultimate limit of "computable" in extended sense
- Boundary between constructive and non-constructive

**As length of hyperarithmetical hierarchy:**

- Measures complexity of second-order definability
- Threshold of  $\Delta^1_1$  comprehension

**As proof-theoretic ceiling:**

- Upper bound for all recursive formal systems
- Horizon of explicit ordinal analysis

**As admissibility threshold:**

- First closure point for constructible hierarchy
- Minimal model of Kripke-Platek

## 8.17 The Gap Between $\Gamma_0$ and $\omega_1^{\text{CK}}$

**What lies between?**

$$\Gamma_0 < \psi_0(\varepsilon_{\Omega+1}) < \psi_0(\varepsilon_{\Omega_2+1}) < \dots < \psi_0(\varepsilon_{I+1}) < \dots < \omega_1^{\text{CK}}$$

where  $I$  = first inaccessible

**Each step** corresponds to:

- Stronger comprehension axioms
- Larger large cardinals reflected
- More complex collapsing functions

Yet all remain **strictly below**  $\omega_1^{\text{CK}}$ .

## 8.18 $\omega_1^{\text{CK}}$ in Contemporary Research

### Computable Analysis

**Effective descriptive set theory** uses  $\omega_1^{\text{CK}}$  as:

- Measure of effective complexity
- Calibration for constructive content
- Boundary of algorithmic methods

### Proof Mining

**Kohlenbach's program** extracts bounds from proofs:

- Proofs in systems with  $|S| < \omega_1^{\text{CK}}$  yield recursive bounds
- Beyond  $\omega_1^{\text{CK}}$ : bounds may be non-constructive

### Reverse Mathematics

**Subsystems of  $\mathbf{Z}_2$**  below  $\Pi^1_1\text{-CA}_0$ :

- All have proof-theoretic ordinals  $< \omega_1^{\text{CK}}$
- $\omega_1^{\text{CK}}$  acts as universal upper bound

## Part 9: Gödel's Dialectica Interpretation

### 9.1 Motivation: Hilbert's Program After Gödel

**Historical context (1958)**: After the incompleteness theorems, Hilbert still hoped to:

1. Extract "finitistic" content from classical proofs
2. Translate classical mathematics into constructive form
3. Obtain computational bounds from existential theorems

**Gödel's idea**: Don't eliminate classical mathematics, but **interpret** it in a more concrete language where:

- Existentials  $\exists x \varphi(x)$  become **functions**  $f$  such that  $\varphi(f(\text{parameters}))$
- Universals  $\forall x$  become **challenges** that functions must overcome

### 9.2 Functional Interpretation: Central Idea

**Basic principle**: Every formula  $\varphi$  is translated into a functional formula  $\varphi^D$  of the form:

$$\exists f \forall x \varphi^*(f, x)$$

where:

- $f$  are **computational witnesses** (functionals)
- $x$  are **counter-witnesses** (challenges)
- $\varphi^*$  is quantifier-free

**Meaning:**

- " $\varphi$  is true" is interpreted as "there exists a program  $f$  that wins against every challenge  $x$ "
- $f$  is a **constructive** object
- $\varphi^*$  is **decidable**

## 9.3 The Dialectica Translation: Formal Definition

For formula  $\varphi$  in the language of Heyting Arithmetic (HA), define  $\varphi^D \equiv \exists f \forall x \varphi^D(f, x, a)$  where  $a$  are free parameters.

### Induction on Structure of $\varphi$

**Atomic formulas:**

$$(t = s)^D \equiv t = s \text{ (no quantifiers added)}$$

**Conjunction:**

$$(\varphi \wedge \psi)^D \equiv \exists f, g \forall x, y [\varphi^D(f, x, a) \wedge \psi^D(g, y, a)]$$

**Disjunction:**

$$(\varphi \vee \psi)^D \equiv \exists z, f, g \forall x, y [(z=0 \rightarrow \varphi^D(f, x, a)) \wedge (z \neq 0 \rightarrow \psi^D(g, y, a))]$$

- $z$  is a "selector" choosing which to prove

**Implication** (crucial case):

$$(\varphi \rightarrow \psi)^D \equiv \exists F, G \forall x, f [\varphi^D(f, Gx, a) \rightarrow \psi^D(Fx, x, a)]$$

- $F$  transforms witnesses of  $\varphi$  into witnesses of  $\psi$
- $G$  produces counter-witnesses for  $\varphi$  from counter-witnesses for  $\psi$

**Negation:**

$$(\neg \varphi)^D \equiv \exists G \forall f [\varphi^D(f, Gf, a) \rightarrow \perp]$$



- G always produces a counter-witness

**Existential quantifier:**

$$(\exists x \varphi(x))^D \equiv \exists t, f \forall y \varphi^D(t, f, y, a)$$

- t is the explicit witness
- f handles counter-witnesses for  $\varphi(t)$

**Universal quantifier:**

$$(\forall x \varphi(x))^D \equiv \exists f \forall x, y \varphi^D(fx, y, a, x)$$

- f is a function providing witnesses for each x

## 9.4 Soundness Theorem

**Theorem (Gödel, 1958):** If  $HA \vdash \varphi$ , then there exists a term t (primitive recursive functional) such that:

$$\vdash \varphi^D(t)$$

**In other words:** Every proof in HA translates into an **explicit algorithm** realizing the dialectica interpretation.

### Consequences

1. **Program extraction:** From proof of  $\exists x \varphi(x)$  we obtain algorithm computing such x
2. **Computational bounds:** From  $\forall n \exists m \psi(n, m)$  we obtain computable function f:  $\psi(n, f(n))$
3. **Relative consistency:** HA is consistent if System T (typed  $\lambda$ -calculus) is consistent

## 9.5 Gödel's System T

To interpret the functionals, Gödel introduced **System T**: a typed  $\lambda$ -calculus with primitive recursion.

### Types

$\mathbb{N}$  (natural numbers)

$\sigma \rightarrow \tau$  (functions from  $\sigma$  to  $\tau$ )

$\sigma_1 \rightarrow \dots \rightarrow \sigma_k \rightarrow \mathbb{N}$  (functionals of finite type)

## Terms

0, S (zero and successor)

$\lambda x.M$  (abstraction)

$M N$  (application)

$R[M,N,P]$  (recursor:  $R[n,z,s] = s^n(z)$ )

## Properties of T

1. **Strong normalization**: every reduction terminates
2. **Confluence**: order of reduction irrelevant
3. **Decidability**: term equality decidable
4. **Expressiveness**: all primitive recursive functions

**Connection**: Functionals extracted from HA proofs are terms in T.

## 9.6 Detailed Example: Mean Value Theorem

**Classical theorem**: For every continuous  $f: [0,1] \rightarrow \mathbb{R}$ ,  $\exists c \in [0,1]: f(c) = \int_0^1 f(x)dx$

### Arithmetical Version

$\forall f \forall n \exists m \forall k < n [|f(m/n) - (\sum_i f(i/n))/n| < 1/k]$

### Dialectica Interpretation

$\varphi^D \equiv \exists F \forall f,n,k [|f(F(f,n,k)/n) - \text{mean}| < 1/k]$

### Extraction

From the proof we obtain explicit algorithm for F:

$F(f, n, k) := \text{first } m < n \text{ such that } |f(m/n) - \text{mean}| < 2/k$

(using Markov principle or modifications)

### Bounds

Complexity of F depends on the proof:

- "Direct" proof:  $O(n \cdot k)$
- "Optimized" proof:  $O(n \cdot \log k)$

## 9.7 Negative Translation: Gödel-Gentzen

To connect classical logic to intuitionistic, we need a **negative translation**.

### Gödel's Translation: $\varphi \mapsto \varphi^N$

$$(t = s)^N \equiv \neg\neg(t = s)$$

$$(\varphi \wedge \psi)^N \equiv \varphi^N \wedge \psi^N$$

$$(\varphi \vee \psi)^N \equiv \neg(\neg\varphi^N \wedge \neg\psi^N)$$

$$(\varphi \rightarrow \psi)^N \equiv \varphi^N \rightarrow \psi^N$$

$$(\exists x \varphi)^N \equiv \neg \forall x \neg \varphi^N$$

$$(\forall x \varphi)^N \equiv \forall x \varphi^N$$

**Theorem:**  $PA \vdash \varphi \Rightarrow HA \vdash \varphi^N$

### Combination

$$PA \vdash \varphi \Rightarrow HA \vdash \varphi^N \Rightarrow \exists \text{ functional } f: (\varphi^N)^D(f)$$

Therefore even from **classical** proofs we extract programs!

## 9.8 Modified Dialectica Interpretation

**Original problem:** The standard interpretation doesn't preserve all logical structures uniformly.

### Solution (Diller-Nahm, Stein): Modified Dialectica

$$(\varphi \rightarrow \psi)^{Dm} \equiv \exists F \forall x [\exists f \varphi^D(f, x) \rightarrow \psi^D(Fx)]$$

### Advantages:

- More modular
- Better bounds
- More direct connection to category theory

## 9.9 Extracting Bounds: Case Studies

### Case 1: Bolzano-Weierstrass Theorem

**Statement:** Every bounded sequence in  $\mathbb{R}$  has a convergent subsequence.

**Computational version:**

$$\forall (a_n), \forall \varepsilon, \exists N, \exists (n_k): \forall k, l > N [|a_{n_k} - a_{n_l}| < \varepsilon]$$

**Dialectica extraction:**

- Functional  $F$ : given  $(a \square)$  and  $\varepsilon$ , produces  $N$  and subsequence
- Complexity:  $F$  uses  $O(2^{\lceil 1/\varepsilon \rceil})$  comparisons
- **Improvement:** With proof analysis,  $O(1/\varepsilon^2)$

**Case 2: Banach Fixed Point Theorem**

**Statement:** Contraction  $f: X \rightarrow X$  has unique fixed point.

**Extraction:**

Number of iterations:  $n \geq \log(d(x_0, x_1)/\varepsilon) / \log(1/c)$

Extracted automatically from classical proof!

**Case 3:  $WKL_0$  (Weak König's Lemma)**

**Statement:** Every infinite binary tree has an infinite branch.

**Dialectica:** Functional that, given oracle for tree, produces sequence constructing the branch.

**Complexity:** The interpretation shows  $WKL_0$  has primitive recursive computational complexity.

**9.10 Connection with Proof-Theoretic Ordinals**

**Theorem (Howard, 1970):** The Dialectica interpretation of PA produces functionals with normalization complexity of ordinal  $\varepsilon_0$ .

**Generalization**

| System                | Ordinal                          | Dialectica Complexity           |
|-----------------------|----------------------------------|---------------------------------|
| PA                    | $\varepsilon_0$                  | Recursion up to $\varepsilon_0$ |
| $ACA_0$               | $\varepsilon_0$                  | Type-1 functionals              |
| $ATR_0$               | $\Gamma_0$                       | Type- $\omega$ functionals      |
| $\Pi^1_1\text{-}CA_0$ | $\psi_0(\varepsilon_{\Omega+1})$ | Bar recursion                   |

**Interpretation:** Complexity of Dialectica interpretation is proportional to proof-theoretic ordinal.

**9.11 Bar Recursion and Extensions**

For systems beyond PA, we need to extend System T with **bar recursion**.

### Spector's Bar Recursion

BR:  $((\mathbb{N} \rightarrow \mathbb{N}) \rightarrow \mathbb{N}) \rightarrow ((\mathbb{N} \rightarrow \mathbb{N}) \rightarrow \mathbb{N} \rightarrow \mathbb{N}) \rightarrow (\mathbb{N} \rightarrow \mathbb{N}) \rightarrow \mathbb{N}$

**Semantics:** Recursion on "bars" (well-founded subsets of  $\mathbb{N}^{<\omega}$ ).

**Use:**

- Necessary to interpret  $\Pi^1_1$ -CA<sub>0</sub>
- Allows extraction from proofs using comprehension
- Complexity: beyond primitive recursion, but still total

**Theorem (Spector, 1962):** System T + bar recursion interprets classical analysis (full  $Z_2$ ) via extended Dialectica.

## 9.12 Proof Mining: Kohlenbach's Program

**Objective:** Extract **quantitative** information from **qualitative** proofs.

### Method

1. Identify fragments of mathematics with good extraction properties
2. Apply Dialectica (or variants)
3. Analyze extracted functionals
4. Obtain explicit computational bounds

### Successes

**Nonlinear functional analysis:**

- Theorem: Successive approximations converge
- Extracted: Explicit convergence rate  $O(1/n)$

**Fixed point theory:**

- Abstract theorems  $\rightarrow$  concrete algorithms with bounds

**Ergodic theory:**

- Ergodic theorems  $\rightarrow$  convergence rates

**Approximation theory:**

- Existence of approximations  $\rightarrow$  constructive algorithms

**Striking results:** Often extracted bounds are **better** than those obtained "by hand"!

## 9.13 Realizability and Dialectica

Comparison with Kleene realizability:

**Kleene:**  $n \Vdash \varphi$  ( $n$  realizes  $\varphi$ )

- Based on partial computations
- $\exists x \varphi$  realized by pair  $(n, m)$ :  $n$  is witness,  $m$  realizes  $\varphi(n)$

**Dialectica:**  $\exists f \forall x \varphi^*(f, x)$

- Total functionals
- Existential-universal symmetry

**Relation**

- Dialectica is "symmetric" realizability
- Dialectica modifications converge toward realizability
- Unified in **dependent choice** frameworks

## 9.14 Categorical Interpretation

**Dialectica as functor:**

There exists functor  $D: \mathbf{HA} \rightarrow \mathbf{T}^{\text{op}} \times \mathbf{T}$  such that:

$$D(\varphi) = (X_{\forall}, X_{\exists}, R: X_{\exists} \times X_{\forall} \rightarrow \Omega)$$

- Objects: types of counter-witnesses and witnesses
- Morphisms: translation of implications

**Properties**

- $D$  preserves conjunctions as products
- $D$  transforms implications into exponentials
- Connection with **Dialectica categories** (de Paiva)

**Modern application:** Category theory provides elegant semantics for program extraction.

## 9.15 Contemporary Variants

### 1. Monotone Dialectica (Kohlenbach)

Restricts to monotone formulas to simplify extraction.

### 2. Bounded Dialectica

For computational complexity: limits type of functionals.

### 3. Dialectica with Side Conditions

Adds side clauses to preserve more structure.

### 4. Dialectica Light

For lightweight programming languages (polynomial time).

## 9.16 Practical Applications

### Formal Verification

- **Coq**: Native extraction of OCaml/Haskell programs from proofs
- **Agda**: Curry-Howard + Dialectica elements
- **Minlog**: Dedicated system for proof mining

### Optimization

- Prove existence of optimal algorithm
- Extract concrete algorithm
- Obtain optimality certificate

### Theoretical Machine Learning

- PAC learning theorems → extracted algorithms
- Generalization bounds → computable via Dialectica

## 9.17 Limitations and Open Problems

### Known Limitations

1. **Type explosion**: Extracted functionals have very high type
2. **Practical complexity**: Though total, may be intractable
3. **Non-compositional**: Combining extractions doesn't always produce optimal extraction

### Open Problems

1. **Optimal characterization**: When does extraction produce optimal bounds?
2. **Automation**: How to fully automate the process?
3. **Generalization**: Extend to ZFC, set theory, ...?
4. **Lower bounds**: When does extraction produce intrinsic complexity?

## 9.18 Philosophy: Hilbert Vindicated?

**Interpretation:** Dialectica interpretation shows that:

1. ✓ **Computational content:** Every proof has extractable algorithmic content
2. ✓ **Extended finitism:** Even classical logic reducible to "finite" objects (functionals)
3. ✗ **Completeness:** But System T requires infinity (high-type functionals)
4. ~ **Hilbert's program:** Partially saved, but in modified form

**Modern consensus:** Dialectica realizes the spirit of Hilbert's program, if not the letter.

## 9.19 Impact on Contemporary Mathematics

### Computer-Assisted Proofs

**Proof assistants** increasingly use Dialectica-like techniques:

- Extract executable code from constructive proofs
- Generate certified algorithms
- Provide computational witnesses automatically

### Complexity Theory

**Implicit Computational Complexity:**

- Characterize complexity classes by logical systems
- Use Dialectica to extract complexity bounds
- Connect proof theory to computational complexity

### Constructive Mathematics

**Bishop-style constructivism** enhanced by:

- Precise extraction methods
- Quantitative information from qualitative theorems
- Computational content made explicit

## Part 10: Predicative Analysis and Subsystems of $Z_2$

### 10.1 Second-Order Arithmetic $Z_2$ : The Complete Framework

**Language  $L_2$ :**

- Number variables:  $n, m, k, \dots$
- Set variables:  $X, Y, Z, \dots$
- Symbols:  $0, S, +, \times, \in, =$



## Axioms of $Z_2$ (Full Second-Order Arithmetic)

1. **Basic axioms of PA** for arithmetic

**Full comprehension:** For every formula  $\varphi(n, X_1, \dots, X_k)$ :

$$\exists Y \forall n [n \in Y \leftrightarrow \varphi(n, X_1, \dots, X_k)]$$

- 2.

**Second-order induction:** For every formula  $\varphi(n, X)$ :

$$[\varphi(0, X) \wedge \forall n (\varphi(n, X) \rightarrow \varphi(n+1, X))] \rightarrow \forall n \varphi(n, X)$$

- 3.

**Power of  $Z_2$ :**

- Standard model:  $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$
- $Z_2 \approx \text{ZFC}$  for "countable" mathematics
- All theorems of classical analysis are provable in  $Z_2$

## 10.2 The Predicativity Question: Poincaré and Weyl

**Impredicative definition:** A definition that quantifies over a class containing the object being defined.

### Classic Example

$$S = \{x \in \mathbb{R} : \forall A \subseteq \mathbb{R} (A \text{ has property } P \rightarrow x \in A)\}$$

Here  $S$  is defined by quantifying over all subsets of  $\mathbb{R}$ , potentially including  $S$  itself!

**Poincaré's objection (1906):** Such definitions are circular and should be prohibited.

**Weyl's program (1918):** Reconstruct mathematical analysis using only predicative definitions.

## 10.3 Ramified Analysis: Russell's Hierarchy

**Idea:** Stratify sets by levels, allowing quantification only over lower levels.

### Level Structure

- **Level 0:** Sets definable from arithmetical formulas (no set quantifiers)
- **Level  $\alpha$ :** Sets definable quantifying only over levels  $< \alpha$
- **Limit level  $\lambda$ :** Union of all levels  $< \lambda$

## Axiomatization $RA(\alpha)$

**Ramified Analysis up to ordinal  $\alpha$ :**

For every  $\beta < \alpha$  and formula  $\varphi$  with quantifiers bounded to levels  $< \beta$ :

$$\exists X^\beta \forall n [n \in X^\beta \leftrightarrow \varphi(n, Y_1^\beta(y_1), \dots, Y_\square^\beta(y_\square))]$$

where  $y_i < \beta$ .

## 10.4 The Feferman-Schütte Theorem

**Fundamental theorem (Feferman-Schütte, 1964):**

1. **Limit of predicativity:**  $RA(\Gamma_0)$  proves exactly the same  $\Pi^1_1$  statements as  $Z_2$
2.  **$\Gamma_0$  is maximal:** For every  $\alpha < \Gamma_0$ ,  $RA(\alpha)$  is strictly weaker
3. **Beyond  $\Gamma_0$  is impredicative:**  $RA(\Gamma_0 + 1)$  introduces genuinely impredicative definitions

**Interpretation:**  $\Gamma_0$  marks the precise boundary between predicative and impredicative.

## 10.5 Detailed Analysis of $ATR_0$

**Arithmetical Transfinite Recursion**

### Axioms

All axioms of  $ACA_0$  plus:

**ATR Schema:** For every arithmetical formula  $\varphi(\alpha, X, Y)$ :

If  $<$  is a well-ordering, then  $\exists F: \forall \alpha [F(\alpha) = \{n : \varphi(\alpha, F \restriction \alpha, n)\}]$

where  $F \restriction \alpha = \{(\beta, m) : \beta < \alpha \wedge m \in F(\beta)\}$

**Proof-theoretic ordinal:**  $|ATR_0| = \Gamma_0$

**Equivalent characterization:**  $ATR_0$  is the maximal predicative system + arithmetical transfinite recursion.

## 10.6 Characteristic Theorems of $ATR_0$

**Theorems equivalent to  $ATR_0$**  (over base  $RCA_0$ ):

## Ordinal Theory

1. **Comparability of well-orderings:** For all well-orderings  $<_1, <_2$ , there exists an isomorphism or embedding
2. **Existence of rank functions:** Every well-ordering has an associated ordinal
3. **Representability:** Every countable ordinal is representable

## Topology

4. **Cantor-Bendixson theorem:** Every Polish space has a perfect derivation until it becomes perfect or empty
5. **Perfect set theorem for closed sets:** Every uncountable closed set contains a perfect subset

## Descriptive Set Theory

6.  **$\Sigma^1_1$ -separation:** If  $A, B$  are disjoint  $\Sigma^1_1$  sets, there exists an arithmetical set  $C$  separating them
7. **Reduction principle for  $\Sigma^1_1$ :** A collection of  $\Sigma^1_1$  sets has a  $\Sigma^1_1$  reduction

## Game Theory

8. **Determinacy of open games:** Gale-Stewart games with open winning set are determined

## Algebra

9. **Ulm's theorem (base form):** Countable abelian groups have well-defined Ulm invariants
10. **Embedding theorems** for countable ordered structures

## 10.7 Explicit Construction: Reaching $\Gamma_0$

**Fundamental sequence for  $\Gamma_0$ :**

Define iteratively:

- $\gamma[0] = 0$

- $\gamma[n+1] = \varphi(\gamma[n], 0)$
- $\Gamma_0 = \sup\{\gamma[n] : n \in \mathbb{N}\}$

### Explicitly

$$\gamma[0] = 0$$

$$\gamma[1] = \varphi(0, 0) = \omega^0 = 1$$

$$\gamma[2] = \varphi(1, 0) = \varepsilon_0$$

$$\gamma[3] = \varphi(\varepsilon_0, 0) = \text{"first fixed point of } \alpha \mapsto \varepsilon_\alpha\text{"}$$

$$\gamma[4] = \varphi(\gamma[3], 0)$$

...

**Crucial property:** Each  $\gamma[n]$  is **autonomously definable**, i.e., definable without reference to  $\gamma[n]$  itself.

## 10.8 The System $\Pi^1_1\text{-CA}_0$ : Beyond the Predicative

### $\Pi^1_1$ -Comprehension:

For every formula of the form  $\forall X \varphi(n, X)$  with  $\varphi$  arithmetical:

$$\exists Y \forall n [n \in Y \leftrightarrow \forall X \varphi(n, X)]$$

**Proof-theoretic ordinal:**  $|\Pi^1_1\text{-CA}_0| = \psi_0(\varepsilon_{\Omega+1})$

This is vastly beyond  $\Gamma_0$ . It requires Bachmann-Howard notation with uncountable ordinals.

## 10.9 Bachmann-Howard Notation for $\Pi^1_1\text{-CA}_0$

### Uncountable ordinals as markers:

- $\Omega$  = first uncountable ordinal
- $\Omega_2$  = second uncountable ordinal
- ...

### Collapsing Function $\psi_\Omega$

Define by induction:

$$\mathbf{C}_0(\alpha, \beta) = \beta \cup \{0, \Omega\}$$

$$\mathbf{C}_{\square+1}(\alpha, \beta) = \mathbf{C}_{\square}(\alpha, \beta) \cup \{\gamma + \delta, \varphi(\gamma, \delta), \psi_\xi(\eta) : \gamma, \delta, \xi, \eta \in \mathbf{C}_{\square}(\alpha, \beta), \xi < \alpha\}$$

$$\mathbf{C}(\alpha, \beta) = \bigcup_{\square} \mathbf{C}_{\square}(\alpha, \beta)$$

$$\psi_\Omega(\beta) = \min\{\gamma < \Omega : \beta \in \mathbf{C}(\alpha, \gamma) \wedge \gamma \notin \mathbf{C}(\alpha, \gamma)\}$$

## Interpretation

- Construct "large" ordinals using  $\Omega$
- $\psi_\alpha$  "collapses" these ordinals to countable ordinals
- The collapsing process generates new ordinals

**Key example:**  $\psi_\Omega(\varepsilon_{\Omega+1}) = \text{ordinal of } \Pi^1_1\text{-CA}_0$

## 10.10 Intermediate Subsystems and Complete Hierarchy

**Strict hierarchy:**

$$\text{RCA}_0 \subsetneq \text{WKL}_0 \subsetneq \text{ACA}_0 \subsetneq \text{ACA} \subsetneq \text{ATR}_0 \subsetneq \text{ATR} \subsetneq \Pi^1_1\text{-CA}_0 \subsetneq \Pi^1_1\text{-CA} \subsetneq \Delta^1_2\text{-CA}_0 \subsetneq \dots \subsetneq \text{Z}_2$$

**Ordinals:**

$$\omega^\omega < \varepsilon_0 < \Gamma_0 < \psi_0(\varepsilon_{\Omega+1}) < \psi_0(\varepsilon_{\Omega_\omega+1}) < \psi_0(\varepsilon_{I+1}) < \dots$$

where  $I$  = first inaccessible.

## 10.11 Model-Theoretic Characterization

**Minimal models:**

| System                | Minimal Model                          | Characterization            |
|-----------------------|----------------------------------------|-----------------------------|
| $\text{RCA}_0$        | $(\mathbb{N}, \text{REC})$             | Recursive sets              |
| $\text{WKL}_0$        | $(\mathbb{N}, \text{REC} \cup \{0'\})$ | + Low basis                 |
| $\text{ACA}_0$        | $(\mathbb{N}, \text{ARITH})$           | Arithmetical hierarchy      |
| $\text{ATR}_0$        | $(\mathbb{N}, \text{HYP})$             | Hyperarithmetical hierarchy |
| $\Pi^1_1\text{-CA}_0$ | $(\mathbb{N}, \Pi^1_1)$                | $\Pi^1_1$ -definable sets   |

**Theorem:** Every model of  $S$  contains the minimal model.

## 10.12 Reverse Mathematics: Striking Results

**Phenomenon:** Many "natural" theorems are equivalent to the Big Five.

**Examples for  $\text{ATR}_0$**

1. **Cantor-Bendixson:** Every Polish space has a Cantor-Bendixson hierarchy

2. **Lusin-Sierpiński**: Every  $\Sigma^1_1$  set is the projection of a  $\Delta^0_2$  set
3. **Perfect set property for  $\Sigma^1_1$** : Every  $\Sigma^1_1$  set is countable or has a perfect subset
4. **Ramsey's theorem** (advanced infinite version)

### Examples for $\Pi^1_1$ -CA<sub>0</sub>

1. **Silver's dichotomy**: For  $\Sigma^1_1$  equivalence, either  $\leq \aleph_1$  classes or  $2^{\aleph_0}$
2. **Borel determinacy** for certain specific games
3. **Embedding theorems** for countable structures
4. **Ulm's theorem** in its most general form

## 10.13 The Boundary of Impredicativity: Why $\Gamma_0$ ?

### Autonomy argument (Feferman):

An ordinal  $\alpha$  is **autonomously accessible** if it can be defined without quantifying over ordinals  $\geq \alpha$ .

### Theorem

- Every  $\alpha < \Gamma_0$  is autonomously accessible
- $\Gamma_0$  is NOT autonomously accessible
- Defining  $\Gamma_0$  requires quantifying over arbitrarily large ordinals  $< \Gamma_0$

### Example

$$\varepsilon_0 = \sup\{\omega, \omega^\omega, \omega^{\omega^\omega}, \dots\}$$

- Definable without using  $\varepsilon_0$

$$\Gamma_0 = \sup\{\varphi(\alpha, 0) : \alpha < \Gamma_0\}$$

- Circular!

To define  $\Gamma_0$ , I must say: "the first fixed point of  $\alpha \mapsto \varphi(\alpha, 0)$ ", which quantifies over all  $\alpha$ , including those  $\geq \Gamma_0$ .

## 10.14 Proof Strength Ordinals: Complete Table

| System | Ordinal | Notation Required |
|--------|---------|-------------------|
|--------|---------|-------------------|

|                               |                                       |                                |
|-------------------------------|---------------------------------------|--------------------------------|
| PRA                           | $\omega^\omega$                       | Finite exponentiation          |
| PA                            | $\epsilon_0$                          | Single fixed points            |
| $ACA_0$                       | $\epsilon_0$                          | Like PA                        |
| $ATR_0$                       | $\Gamma_0$                            | Veblen + fixed point $\Gamma$  |
| $\Pi^1_1\text{-}CA_0$         | $\psi_0(\epsilon_-(\Omega+1))$        | Bachmann-Howard, $\Omega$      |
| $\Delta^1_2\text{-}CA_0 + BI$ | $\psi_0(\epsilon_-(\Omega_\omega+1))$ | Extension with $\Omega_\omega$ |
| $\Pi^1_2\text{-}CA_0$         | $\psi_0(\epsilon_-(M+1))$             | Mahlo $M$                      |
| KP                            | $\psi_0(\Omega_\Omega)$               | Inaccessible                   |
| KPi                           | $\psi_0(K)$                           | Weakly compact $K$             |

## 10.15 Applications of the Predicative Frontier

### 1. Foundations of Mathematics

- $\Gamma_0$  separates "iterative constructible" from "essentially impredicative"
- Philosophical debate: Is  $\Gamma_0$  the limit of "safe" mathematics?

### 2. Computer Science

**Proof objects**  $< \Gamma_0$ :

- Languages like Coq: proof objects  $< \Gamma_0$
- Beyond  $\Gamma_0$ : Requires reflection, universes

### 3. Analysis

**Classical undergraduate analysis**: Everything in  $ATR_0$

**Modern measure theory**: Requires  $\Pi^1_1\text{-}CA_0$

### 4. Algebra

**Basic group theory**:  $ACA_0$

**Infinite group theory**:  $ATR_0, \Pi^1_1\text{-}CA_0$

## 10.16 Open Problems at the Frontier

1. **Natural weakening of  $ATR_0$** : Is there a natural system between  $ACA_0$  and  $ATR_0$ ?

2. **Ordinal of full  $Z_2$** : We only know bounds, not precise value
3. **Ultimate predicative levels**: Is there interesting structure between  $\Gamma_0$  and  $\psi_0(\varepsilon_{\Omega+1})$ ?
4. **Homeless theorems**: Few natural theorems lie strictly between Big Five—why?
5. **Higher-order reverse mathematics**: What happens for  $Z_3, Z_4, \dots$ ?

## 10.17 The Predicative vs Impredicative Divide

**Predicative mathematics** ( $\leq \Gamma_0$ ):

- Definitions refer only to previously constructed objects
- "Bottom-up" construction
- Computationally meaningful
- Philosophically conservative

**Impredicative mathematics** ( $> \Gamma_0$ ):

- Definitions can quantify over collections including defined object
- "Top-down" specification
- May lack computational content
- Philosophically bold

### Contemporary View

**Consensus**: The predicative/impredicative distinction is:

- Mathematically precise ( $\Gamma_0$ )
- Philosophically significant
- Computationally relevant
- But not an absolute barrier—both are legitimate

## 10.18 Subsystems and Mathematical Practice

**Observation**: Most "ordinary" mathematics fits in surprisingly weak systems.

### Examples by System

**$RCA_0$** : Basic calculus, elementary analysis

**$WKL_0$** : Intermediate value theorem, Heine-Borel

**$ACA_0$** : Bolzano-Weierstrass, countable choice

**$ATR_0$** : Cantor-Bendixson, comparability



$\Pi^1_1$ -CA<sub>0</sub>: Perfect set theorem, uniformization

**Implication:** Advanced set-theoretic methods are rarely needed for concrete mathematics.

## 10.19 Philosophical Implications

The ordinal hierarchy reveals:

1. **Natural stratification:** Mathematics has intrinsic levels of complexity
2. **Predicativity is precise:** Not a vague philosophical notion but mathematically exact ( $\Gamma_0$ )
3. **Strength is measurable:** Logical strength has an absolute scale (ordinals)
4. **Trade-offs:** Weaker systems = less power but more constructive content
5. **Universality:** The same ordinals appear across logic, computation, set theory